
(MT)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPsrch nn n.a. - n.a. database search, using Smith-Waterman algorithm

```
Run on: Thu Jul 9 15:13:00 1998; MasPar time 4882.54 Seconds
1521.186 Million cell updates/sec
```

Tabular output not generated.

Title: >US-08-717-294-42

Description: (1-4451) from US08717294.seq

Perfect Score: 4451

N.A. Sequence: 1 AAGCTTAACCATGCCCCATG.....ACCTGTACTGAAGCGGGCCGC 4451

Comp: TTCGAATTGGTACGGGTAC.....TGGACATGACTTCGCCGGCG

Scoring table: TABLE default

Gap 6

Nmatch STD : Dbase 0: Query 0

Searched: 457396 seqs, 834335066 bases x 2

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database: emb154

1:em ba 2:em hta 3:em huml 4:em hum2 5:em in 6:em om

7:em_ or 8

Database: aenbank106

13:qb ba 14:qb hta 15:qb in 16:qb om 17:qb ov 18:qb pat

19:qb ph 20:qb pl 21:qb pr1 22:qb pr2 23:qb ro 24:qb st

25:qb sv 26:qb un 27:qb vi

Statistics: Mean 12.170; Variance 6.902; scale 1.763

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result	No.	Score	Match	Query	ID	Description	Pred. No.
1	1612	36.2	4278	18	108644	Sequence 4 from Patent	0.00e+00
2	1610	36.2	4281	18	108643	Sequence 3 from Patent	0.00e+00
3	1087	24.4	6999	18	171105	Sequence 1 from Patent	0.00e+00
4	1085	24.4	7212	18	105404	Sequence 3 from Patent	0.00e+00
5	1085	24.4	8241	18	A05328	Synthetic Factor VIII	0.00e+00
6	1085	24.4	8241	18	A07042	Artificial mRNA for fo	0.00e+00

	7	1085	24.4	8831	18	E00527	Complete cDNA encoding	0.00e+00
	8	1085	24.4	9029	21	HUMFVIIIC	Human coagulation factor	0.00e+00
	9	1083	24.3	7056	18	I270663	Sequence 1 from patent	0.00e+00
	10	1081	24.3,	7440	18	108345	Sequence 2 from Patent	0.00e+00
	11	1083	24.3	8967	21	HSFVIIIR	Human mRNA for factor	0.00e+00
	12	1083	24.3	8967	18	171409	Sequence 1 from patent	0.00e+00
	13	1083	24.3	9009	18	131901	Sequence 3 from patent	0.00e+00
	14	1083	24.3	9009	21	HUMFVIII	Human coagulation factor	0.00e+00
	15	1083	24.3	9009	18	163424	Sequence 1 from patent	0.00e+00
	16	1073	24.16	7440	18	108457	Sequence 1 from Patent	0.00e+00
	17	1060	23.8	4548	18	108642	Sequence 2 from Patent	0.00e+00
	18	1061	23.8	4551	18	108641	Sequence 1 from Patent	0.00e+00
	19	1020	22.9	6339	16	SFO49517	Us scrofa factor VIII	0.00e+00
	20	1001	22.5	7032	16	AAU16234	Canis familiaris facto	0.00e+00
	21	991	22.3	7145	16	AF049489	Canis familiaris facto	0.00e+00
	22	826	18.6	7493	18	163427	Sequence 7 from patent	0.00e+00
	23	826	18.6	7493	23	MUSCVCIII	Mus domesticus coagula	0.00e+00
	24	758	17.0	1728	18	102054	Sequence 8 from Patent	0.00e+00
	25	664	14.9	1623	18	102064	Sequence 3 from patent	0.00e+00
	26	606	13.6	1993	18	102047	Sequence 1 from Patent	0.00e+00
	27	496	11.1	3852	18	104400	Sequence 35' from Paten	0.00e+00
	28	495	11.1	3852	18	E00422	Human genomic DNA codi	0.00e+00
	29	487	10.9	1130	18	131900	Sequence 1 from patent	0.00e+00
	30	487	10.9	1130	18	163426	Sequence 5 from patent	0.00e+00
	31	339	7.6	2493	21	HUMF8C	Homo sapiens coagulati	1.62e-216
	32	310	7.0	955	18	E00526	Part of pESVDA vector	4.20e-195
	33	242	5.4	1728	18	102049	Sequence 3 from Patent	3.13e-145
	34	142	3.2	5031	21	HUMFAC15	H.sapiens factor VIII	2.22e-73
	35	129	2.9	1128	21	HUMFAC08	H.sapiens factor VIII	2.83e-64
	36	123	2.8	391	21	HUMFAC12	H.sapiens factor VIII	4.21e-60
	37	123	2.8	1728	18	102052	Sequence 6 from Patent	4.21e-60
	38	111	2.5	293	21	HUMFAC20	H.sapiens factor VIII	8.05e-52
	39	112	2.5	669	18	104388	Sequence 23 from Paten	1.66e-52
	40	113	2.5	4074	21	HUMFAC13	H.sapiens factor VIII	3.41e-53
	41	112	2.5	15155	18	104391	Sequence 26 from Paten	1.66e-52
	42	101	2.3	346	21	HUMFAC07	H.sapiens factor VIII	5.36e-45
	43	102	2.3	2709	21	HUMFAC21	H.sapiens factor VIII	1.12e-45
	44	96	2.2	218	21	HUMFAC10	H.sapiens factor VIII	1.29e-41
	45	96	2.2	367	21	HUMFAC17	H.sapiens factor VIII	1.29e-41

ALIGNMENTS

1	RESULT	108644	4278 bp	PAT	14-NOV-1994
	LOCUS	Sequence 4 from Patent WO 8800831.			
	DEFINITION	108644			
	ACCESSION	g588650			
	NID	.			
	KEYWORDS				
	SOURCE	Unknown.			
	ORGANISM	Unknown.			
	REFERENCE	Unclassified.			
	AUTHORS	1 (bases 1 to 4278)			
	TITLE	Pasek, M.P.			
		DNA SEQUENCES CODING FOR MODIFIED FACTOR VIII:C AND MODIFIED FACTOR			
		VIII:C-LIKE POLYPEPTIDES AND PROCESSES FOR PRODUCING THESE			
		POLYPEPTIDES IN HIGH YIELDS			
	JOURNAL	Patent: WO 8800831-A 4 11-FEB-1988;			
	FEATURES	Location/Qualifiers			
	source	1..4278			
		/organism="unknown"			
	BASE COUNT	1244 a	942 c	946 g	1146 t
	ORIGIN				

Matches 3211; Conservative 0; Mismatches 1061; Indels 90; Gaps 26;

Db 4 GCCACCAAGAGATACCTACTGCTGGTGCAGTGAACCTGTCATGGGACTATATGCAAGTGT 63
 Qy 84 GCCACCGCGCTACTACTGCTGGGCGGTGAGCTGCTGGGACTATATGCAAGTGT 143
 Db 64 CTCGGTGAAGTCTGTCGAGCGCAAGATTTCTCTGAGTGGCAAAATCTTTTCCATT 123
 Qy 144 CTGGGCGAGCTCCCGTGGAGCGCGCTTCCCGCCCGCTGCGCCAGAGCTTCCCGCTC 203
 Db 124 AAGACCTCAGCTGCTACAAAAGAGACTGCTTTGTAGAAATTCACGATCACCCTTTTCAAC 183
 Qy 204 AACACCAAGCTGGTGTACAGAAACCTGCTGCTGGAGTTTACACGACCTGTTTCAAC 263
 Db 184 ATCGCTAAGCGCAAGCGCCACCTGAGTGGGTCTGCTAGCTCTACCATCGAGCTGAGGTT 243
 Qy 264 ATTGCCAAGCGCGCCCGCTGCTGATGGGCTGCTGGGCGCCACCATCCAGGCGGAGTG 323
 Db 244 TATGATACAGTGGTCACTTACACTTAAGACATGGCTTCCCATCTCTGTCATGCTTCTATGCT 303
 Qy 324 TAGGACACCTGCTGATCACCCTGAAGACATGGCCAGCCACCCGCTCAGCTGCACGCC 383
 Db 304 GTTGTGTATCTCTGCAAGCTTCTGAGGAGCTGAATATGATGATGACACCACTCAA 363
 Qy 384 CTGGGCTGAGCTACTGGAAGCGCAGCGGGCGGAGTACCGACGACGACCTCCAG 443
 Db 364 AGGGACGAAGAGATGATAAGTCTTCCCTGGTGAAGCCATACATATGCTGGCAGGTC 423
 Qy 444 CGCAGAGAGGAGGACGACAAAGTGTTCGGGGGGGAGCCACACCTAGCTGGCAGGTC 503
 Db 424 CTGAAGAGATGGTCCATGGCTCTGACCCACTGTCGCTTACCTACTCATATCTTTCT 483
 Qy 504 CTTAAGGACAGCGCCCTATGGCCAGCCCGCTCTGCTGACCTACAGCTACCTGAGC 563
 Db 484 CATGTGACCTGGTAAAGACTTGAATTCAGGCTCATTTGAGCCCTACTAGTATGATA 543
 Qy 564 CAGCTGACCTGGTGAAGGATCTGAACAGGGGCTGATCGGCGGCTGCTGCTGCTG 623
 Db 544 GAAGGAGTCTGGCCAGAAAGACACACTTGCACAAATTTATATCTACTTTTGTCT 603
 Qy 624 GAGGCGAGCTGGCCAGGAGAAACCCAGCCCTGCACAGTTTATCTCTGCTGCTGCTG 683
 Db 604 GTATTTGATGAAGGAAAGTGGCACTCAGAAACAAAGAACTCTTGTGACGAGATAGG 663
 Qy 684 GTGTTTCAGAGGGAAGAGCTGGCAGACGCGAGCTAAGAACAGCTGATCGAGGACCG 743
 Db 664 GATGCTCATCTCTCGGCGCTGGCTTAAATTCACACAGCTCAATGGTTATGTAAACAGG 723
 Qy 744 GACGCGCCAGCGCGCGCTGGCCCAAGATGCACCCGTTAACGGCTACGCTGAACCGC 803
 Db 724 TCTCTGATACAGGCTGATGATGATGCCACAGAAATCAGTCTTATGGCATGCTGATGA 783
 Qy 804 AGCCTGC--CC--GGCTGATGGCTGGCCAGGAGCGCTGATGGCAGCTCATCGGC 860
 Db 784 ATGGGCAACACTCTGAAAGTGCATCAATATTCCTGCAAGGTCAACATTTCTTTGTGAGG 843
 Qy 861 ATGGGCAACACTCTGAAAGTGCATCAATATTCCTGCAAGGTCAACATTTCTTTGTGAGG 920
 Db 844 AACCATCGGCGGCTCTTGGAAATCTCGCAATTAACCTTTCTTACTGCTCAAACTC 903
 Qy 921 AACCCAGCGGCGGCGGCTGGAGATCAGCCCGCATCACCCTTCTGACTGCCCGACCTG 980
 Db 904 TTGATGACCTTGGACAGTTTCTACTGCTTTTGTGCTATCTCTTCCACCAACATGATGGC 963

Db 3697 ATTTTAACCTCCAAATTAATGCTGATACATCGTTTGCACCAACATCATTATGACATT 3756
 Qy 3858 ATCTTCAACCCCCCATCATCGCCGCTACATCGGCTGCACCCACCCATCAGCATC 3917
 Db 3757 CGCAGCACTCTTCGATGAGTGTGATGGCTGTGATTTAAATAGTTGCAGCATGCCATTG 3816
 Qy 3918 CGCAGCACTTCGATGAGTGTGATGGCTGTGATTTAAATAGTTGCAGCATGCCATTG 3977
 Db 3817 GGAATGAGAGTAAAGCAATATCAGATGACAGATTAATGCTTCTGCTTACTTACCAAT 3876
 Qy 3978 GGCATGGAGAGCAAGGCCATCAGCGCGCCAGATCAGCGCTCCAGCTTACTTACCAAC 4037
 Db 3877 ATGTTTGCACCTGCTCTTCTTCAAGCTCGACTTCACTCCAAAGGAGGATTAATGCC 3936
 Qy 4038 ATGTTTGCACCTGCTCTTCTTCAAGCTCGACTTCACTCCAAAGGAGGATTAATGCC 4097
 Db 3937 TGGAGACCTCAGGTGAATATCAAAAGAGTGGTGCAGTGGACTTCCAGAGACAAATG 3996
 Qy 4098 TGGCGCCCGCAGGTGAACACCCAGAGTGGCTGAGTGGACTTCCAGAGACAAATG 4157
 Db 3997 AAAGTCAAGGAGTAACTACTCAGGAGTAAATCTCTCTTACAGCATGTGCTGAAG 4056
 Qy 4158 AAGGTGACTGGGCTGACCAACCCAGGCGCTCAGAGCTCTGACCAAGATGTGAGTGAAG 4217
 Db 4057 GAGTCTCTATCTCAGAGTCAAGATGCCATCAGTGGACTCTCTTTTTCAGATGCC 4116
 Qy 4218 GAGTCTCTATCAGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 4277
 Db 4117 AAGTAAAGGTTTTTCAGGGAATCAAGACTCTTCACTCTGCTGCTGCTGCTGCTGCTG 4176
 Qy 4278 AAGTAAAGGTTTTTCAGGGAATCAAGACTCTTCACTCTGCTGCTGCTGCTGCTGCTG 4337
 Db 4177 CCAGCTTACTGCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCT 4236
 Qy 4338 CCGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4397
 Db 4237 CTGAGGTGAGGCTTCTGGGCTGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 4278
 Qy 4398 CTGCGCATGAGGCTGCTGGGCTGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 4439

RESULT 2
 LOCUS 108643 4281 bp PAT 14-NOV-1994
 DEFINITION Sequence 3 from Patent WO 8800831.
 ACCESSION 108643
 NID g588649
 KEYWORDS .
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 4281)
 AUTHORS Pasek,M.P.
 TITLE DNA SEQUENCES CODING FOR MODIFIED FACTOR VIII:C AND MODIFIED FACTOR VIII:C-LIKE POLYPEPTIDES AND PROCESSES FOR PRODUCING THESE POLYPEPTIDES IN HIGH YIELDS
 JOURNAL Patent: WO 8800831-A 3 11-FEB-1988;
 FEATURES Location/Qualifiers
 source 1..4281
 /organism="unknown"
 BASE COUNT 1245 a 943 c 946 g 1147 t
 ORIGIN
 Query Match 36.2%; Score 1610; DB 18; Length 4281;
 Best Local Similarity 73.6%; Pred. No. 0.00e+00;

Jul 9 17:02

US-08-717-294-42.rge

11

Qy 2901 CAGCACCATGGCCGCCACCAGGACGAGTTGACCTGCAAGGCGCTGGGCGCTACTTCAGC 2960

Db 2800 GATGTTGACCTGGGAAAGATGTGCACTCAGCGCTGATTTGAGACCCCTTCTGGTCTGGCAC 2859

Qy 2961 GAGCTGGACCTGGAAGAAGCGTGCAACAGCGGCGCTGATCGGCGCCCTGCTGGTGTGGCAC 3020

Db 2860 ACTAACCACTGAAACCGCTGCTCATGGGAGCAACAGTGACAGTAAGAAATTTGCTGCTTT 2919

Qy 3021 ACCACACCTTGAAACCCGCCAGGAGGAGCGAGTGACTGTGCGAGGAATTTGCCGTGTTCC 3080

Db 2920 TTCTCCACCATCTTTGATGAGACAAAGCTGTGTACTTCACTGAAATATGGAAGAAAC 2979

Qy 3081 TTC---ACCATCTTGAAGAGACTAAGAGCTGTGTACTTCACGAGAACTAGGAGCGCAAC 3137

Db 2980 TGCAGGGCTCCCTGCAATATCCAGATGGAGATCCACATTTTAAAGAGAAATATCGGCTTC 3039

Qy 3138 TGGCGCGCCCTCGCAACATCCAGATGGAAGATCCACCTTCAAGGAGAACTACCGCTTC 3197

Db 3040 CATGCAATCAATGGCTACATATGGAATACACTACTGCGCTTAGTAATGGCTCAGGATCAA 3099

Qy 3198 CAGGCCATCAACGCGTACATCATGGACACCTCGCGGCGTGTGGTGGCCGAGGACCA 3257

Db 3100 AGSATTGCGTGGTATCTGCTCAGCATGGGAGCAATGAAACATCCATTTATTATTC 3159

Qy 3258 CGCATCGCGTGGTACCTGCTGTATGGGAGAGAACGAGACATCCACAGCATCCACTTC 3317

Db 3160 AGTGAGCATGTGTTCACTGTACGAAAAAAGAGGATATATAATGGCACTGTACAACTC 3219

Qy 3318 AGCGGCGACGTTTTCACGCTGGGCAAGAGAGGAGTACAGATGGCGCTGTACAACTG 3377

Db 3220 TATCCAGGTGTTTTTACAGAGTGGAATGTATACCATCCAAAGCTGGAATTTGGCGGGTG 3279

Qy 3378 TACCCGCGCGTGTTCGAGCATGTGAGATGTGCCAGCAAGCGCGGATCTGCGCGGTG 3437

Db 3280 GATAGCTTATTCGGCAGCATCTACATGCTGGGATGAGCACATTTTCTGGTGTACAGC 3339

Qy 3438 GAGTGCCTGATCGCGCAGCACTGCACGCGGCGATGAGCACCTGTTCTGCTGTGTACAGC 3497

Db 3340 AATAAGTCTCAGACTCCCTGGGAAATGGCTTCTGGACACATTAGAGATTTTCAGATTACA 3399

Qy 3498 AACAAGTCCGAGACCCCTGGGATGGCGAGCGGCGCACATCCGCACTTCAGATCAC 3557

Db 3400 GCTTCAGGACAAATATGAGACTGGGCGCCCAAGCTGGCCAGACTTCATTATTTCGGATCA 3459

Qy 3558 GCGAGCGGCGAGTACGGCGAGTGGGCTCCCAAGCTGGCCCGGCTGCACTACAGCGGAGC 3617

Db 3460 ATCAATGCTGGAGCAGCAAGGCGCTTTTCTTGGATCAAGGTGGATCTGTTGGCACCA 3519

Qy 3618 ATCAAGCGCTGGTGCAGCAAGGAGCGCTTCTCTGGATCAAGGTGGACCTGCTGGCGCCC 3677

Db 3520 ATGATTATTCAGCGGATCAAGACCGAGGTGCCCGTCAGAGTTCTCCAGCTCTACATC 3579

Qy 3678 ATGATCATCCAGCGCATCAAGCCAGGCGCGCGGCGAGACTTCAGCAGCCTGTACATC 3737

Db 3580 TCTCAGTTTATCATATGATGTAGTCTTGTGGAATGTGGATTCATCTGGATATAACAAAT 3639

Qy 3738 AGCCAGTTTATCATATGATGTACTCTTAGACGGGAGAGTGGGAGACTACCGCGGCAAC 3797

Db 3640 TCCAGTGGACCTTAATGGCTTCTTGGCAATGTGGATTCATCTGGATATAACAAAT 3699

Qy 3798 AGCACCGGACCCCTGATGGTGTCTTCGGGCAACGTGGACAGCGCGCATCAAGCAAC 3857

Db 3700 ATTTTAAACCTCAATATTATGCTCGATACATCGTTTTCACCCAACTCATTTAGATT 3759

Jul 9 17:02

US-08-717-294-42.rge

12

Qy 3858 ATCTTCAACCCGCCCATCATCGCGCGCTACATCGCGCTGCACCCACCCACTACAGCATC 3917

Db 3760 CCGAGCACTCTTCGATGGAGTGTATGGGCTGTGATTTAAATATGTTGAGCATGCCATTG 3819

Qy 3918 CCGAGCACTCTTCGATGGAGTGTATGGGCTGTGATTTGACAGCTGACAGCTGCGCCCTG 3977

Db 3820 GGATGGAGAGTAAGCAATATCATGATGCACAGATTTACTGTTTCATCTTACTTTACCAAT 3879

Qy 3978 GCGATGGAGAGAGGCAATCATGACGACGCCCAATCATCGCGCTCGAGCTACTTCAACCAAC 4037

Db 3880 ATGTTTGCACCTGTCTCTTCAAAAAGCTCGACTTTCACCTGCCAAGGGAGGATTAATGCC 3939

Qy 4038 ATGTTGCCACCTGAGCGCCAGCAAGGCGCGCTGCACTTGCAGGGCGCGCAACAGCC 4097

Db 3940 TGGAGACCTCAGGTCATATATCCAAAAGTGGCTGCAAGTGGACTTTCAGAGACAAATG 3999

Qy 4098 TGGCGCCCGCAGGTCAACAAACCCCAAGGAGTGGCTGCAGGTGGACTTCCAGAAAACCATG 4157

Db 4000 AAGTTCACAGGAGTACTACTCAGCGAGTAAATCTCTGCTTACAGCATGTATGTGAAG 4059

Qy 4158 AAGTGTACTGGCTGCACCCAGCGGCTCAAGAGCGCTGTGACAGCATGTAGCTGAAG 4217

Db 4060 GAGTTCCTCATCTCCAGCAGTCAAGATGGCCTCATGAGCTCTCTTTTTCAGAAATGGC 4119

Qy 4218 GAGTTCCTCATCAGCAGCAGCAGCGCCACCACTGAGTGGACCTGTTCTTCCAAAAGCGC 4277

Db 4120 AAGTAAAGGTTTTTTCAGGGAATCAAGACTCTCTTACACCTGTGTGAACCTCTCTAGAC 4179

Qy 4278 AAGTGAAGGTGTTTCAGGGCAACCCAGACAGCTTTCACACCGGTCTGCAACAGCTGCAC 4337

Db 4180 CACCGCTTACTGACTCTGCTACTCTCGAATTCACCCCGCAGAGTTGGGTGCACCAATGGC 4239

Qy 4338 CCGCCCTGCTGACCGCTACCTGCGCATCCACCCAGAGCTGGGTGCACCAATCGCC 4397

Db 4240 CTGAGGATGGAGTTCTGGGCTGCAGGCAAGGACCTCTTAC 4281

Qy 4398 CTGCGCATGAGGTGCTGGGCTGCGAGGCCCGACGACCTGTAC 4439

RESULT 3

LOCUS I71705 6999 bp DNA PAT 23-DEC-1997

DEFINITION Sequence 1 from patent US 5693499.

ACCESSION I71705

NID g3013259

KEYWORDS .

SOURCE Unknown.

ORGANISM Unknown.

REFERENCE 1 (bases 1 to 6999)

AUTHORS Yonemura,H., Tajima,Y., Sugawara,K. and Masuda,K.

TITLE Process for preparing human coagulation factor VIII protein complex

JOURNAL Patent: US 5693499-A 1 02-DEC-1997;

FEATURES Location/Qualifiers

source 1. .6999

BASE COUNT 2224 a 1503 c 1480 g 1792 t

ORIGIN

Query Match 24.4%; Score 1087; DB 18; Length 6999;

Best Local Similarity 76.1%; Pred. No. 0.00e+00;

Matches 1585; Conservative 0; Mismatches 498; Indels 0; Gaps 0;

Db 4917 CCCACGACTGTGAACGCCATCAACGGGAATACTCGTACTACTCTTTCAGTCACATCA 4976

Jul 9 17:02

US-08-717-294-42.rge

13

QY 2360 CCCCCCGCTGTAAGCGCCACCAGCGGAGATCACCCGACCACCTCGMAAGGACCA 2419

Db 4977 AGAGGAATTCAGTATGATGATACCATATACAGTTGAAATGAGAGGAAGATTTTGACAT 5036

QY 2470 GGAGGAGATCGACTAGGACGACCATCAGGCTGGAGATGAGAGGAGGACTTCGACAT 2479

Db 5037 TTATGATCAGCATGAAATACAGACCCCGCAGCTTTCAAAAGAAAACACGACATATT 5096

QY 2480 CTACGACGAGGACGAGACGAGACCCCGCTCTCTCCAAAGAAACCCGACACTATT 2539

Db 5097 TATTGCTGAGTGGAGGCTCTGGGATTTATGGATGAGTAGCTCCGCCACATCTTCAAG 5156

QY 2540 CATCGCGCGCTGGAGCGCTGTGGGACTAGCGCATGAGCAGGACCGCCAGCTCTGCG 2599

Db 5157 AAGAGGGCTCAGAGTGGCAGTGTCCCTCAGTTCAAGAAAGTTGTTTTCAGGAATTTAC 5216

QY 2600 CAACCGCGCCAGAGCGGAGCGCTGCCAGTTCAAGAGGTGCTTTCCAGGAGTTCA 2659

Db 5217 TGATGGCTCTTTACTCAGCCCTTTATACGGTGGAGAACTAAATGACATTTGGGACTCT 5276

QY 2660 CGAGCGCAGCTTCAACGAGCCCTGTACCGCGGAGCTGAACGAGCACTGSGGCTGCT 2719

Db 5277 GGGGCCATATATAGACAGAAAGTTGAGATATATCATGCTAACTTTTCAGAAATCAGCG 5336

QY 2720 CGGCGCTTACATCCGCGGAGGTGAGAGACATCATGTGTGACCTTCGCGACCAAGC 2779

Db 5337 CTCTGCTCCCTATTCTCTATTCTAGCCCTATTCTTATGAGGAGATCAGAGGCAAGG 5396

QY 2780 CTCGCGGCGCTACTCTCTTACTCTCTCCCTGATCAGCTACGAGGAGGACGAGCGCAGG 2839

Db 5397 AGCAGACCTAGAAATAACTTTGTCAAGCCTTAATGAACCAAACCTTACTTTTGGAACT 5456

QY 2840 CGCGGAGCGCGCAAGAACTTCGTGAAGCCCAACGAGACTAAGACCTTCTTGGAACT 2899

Db 5457 GCACATCATATGGACCGCACTAAGATGATTTTGACTGCAAGCCCTGGGCTATTTCCT 5516

QY 2900 CGACACACATGGCCCCACCAAGGACGAGTTGCACTGCAAGSCCTGGGCTTACTTCAG 2959

Db 5517 TGATGTCAGCTGGAAAGAATGTGCACTAGGCTGATTTGGACCCCTTCTGCTGCGCA 5576

QY 2960 CGAGCTGACCTCGGAGGAGCTGCACAGCGGCTGATCGGCCCCCTGCTGGTGGCA 3019

Db 5577 CACTAACACACTGAACCTGCTCATGGGAGACAAAGTGACAGAAATTTGCTCTGTT 5636

QY 3020 CACCAACACCTGAACCCCCCAGCGGAGGCGTGACTGTGAGGAATTTGCCCTGTT 3079

Db 5637 TTTTACCATCTTTGATGAGACCAAAAGCTGTACTTCACTGAAATATGAAAGAAACTG 5696

QY 3080 CTTCACCATCTTCAGAGACTAAGAGTGTGTACTTTTCCAGGAACATGGAGCGCACTG 3139

Db 5697 CAGGCTCCCTGCAATATCGAGATGGAATCCCACTTTTAAAGAGAAATATTCGTTCCA 5756

QY 3140 CCGCGCCCCGCAACATCCAGATGGAAGTCCCACTTCAAGGAGAACTTACCGTTCCA 3199

Db 5757 TGAATCAATGGCTACATATGATGATACCTACCTGGCTTAGTAATGGCTAGGATCAAG 5816

QY 3200 CGGCATCAAGGGCTACATGACACCTTCGCGGCTGGTGTATGGCCAGGACGAGG 3259

Db 5817 GATTGATGTTATCTGCTCAGATGGCAGCAATGAACATCCAATTCATTCAATTCAG 5876

QY 3260 CATCGCTGGTACTGCTGCTATGGGAGGACGAGAAACATCCAAGCATCCACTCAG 3319

Db 5877 TGGCATGTTCTACTGTACGAABAAAGAGGACTATAAATGGCACTGTGCAATCTCTA 5936

Jul 9 17:02

US-08-717-294-42.rge

14

QY 3320 CGGCGCAGCTTTTACCCTGCGCAAGAGGAGGTACAAGATGGCCCTGTACACCTGTA 3379

Db 5937 TCCAGGTGTTTTTGACAGACTGGAAATGTTACCATCCAAAGCTGCAATTTGGCGGTGA 5996

QY 3380 CCGCGGCTGTTTCGAGACTGTGAGATCTGCCCCAGGAGCGCGGATCTGGCGGTGA 3439

Db 5997 ATCCGCTTATTTGGCAGCATCATATGCTGGGATGAGACACATTTTCTGCTGTACAGCAA 6056

QY 3440 GTCCCTGATCGCGGAGACACCTGCACCGCGCATGAGACCCCTGTTCTTGGTGTACAGAA 3499

Db 6057 TAAAGTGCAGACTCCCTGGGATGGCTTCTGCACATTAAGAGATTTTCAGATTACAGC 6116

QY 3500 CAAAGTGCACACCGCCCTGGGCATGGCGAGCGGCCACATCCGCGACTTCCAGATCACCGC 3559

Db 6117 TTCAGGCAATATGACAGCTGGGCCCCAAAGCTGGCGAGACTTCATTATTCGCGATCAAT 6176

QY 3560 CAGCGGCGACTAGCGCCAGTGGGCTCCAAAGCTGGCGGCTGCACTACAGCGGCGAGCAT 3619

Db 6177 CAATGCCGTGACGAGCAAGGAGCCCTTTTCTTGGATCAAGGTGGATCTGTTTGGGCAACAT 6236

QY 3620 CAAAGCTGCTGACGAGGAGGCGCTTCTCTGATCAAGGTGGAGCTGCTGGCCCCCAT 3679

Db 6237 GATATTCAGGGCATCAAGACCCAGGCTGCCGTCAGAAAGTTCTCCAGCCTCTACATCTC 6296

QY 3680 GATCATCCAGGCAATCAAGACCCAGGCGGCCCGCAGAGTTCCAGCGCCTGTACATCAG 3739

Db 6297 TCAGTTTATCATCATGTATAGTCTTGTGGAAGAGTGGCGAGCTTATTCAGGAAATTC 6356

QY 3740 CCAGTTTATCATCATGTACTCTCTAGACGGCAAGAGTGGCAGACTTACCGCGGCAACAG 3799

Db 6357 CACTGGAACTTAATGGTCTTTTGGCAATGTGATTCATCTGGGATTAAGATTAACATAT 6416

QY 3800 CACCGGCAACCTGATGCTGTTCTTGGCAAGTGGCAGCGGCAATCAAGCAACAT 3859

Db 6417 TTTTAACCTTCAATATATGCTGATATCATCGTTTGCACCACTCAATATATAGATTCG 6476

QY 3860 CTTCACCCCCCATCATCGCGGCTAGATCGGCTGCACCCCACTACAGCATCCG 3919

Db 6477 CAGCACTTTCGCTGAGTGTGATGGCTGTGATTTAAATAGTTGCAGCATGGCAATGGG 6536

QY 3920 CAGCACCTTCGCGCATGAGCTGATGGGCTCGACCTGACAGCTGCAGCATGGCCCTGGG 3979

Db 6537 AATGAGAGTAAAGCAATATCAGATGCACAGATTACTGCTTCACTACTTTTACCAATAT 6596

QY 3980 CATGAGAGCAAGGCGCATCAGCGCCGCGAGATCAGCGCTTCCAGCTACTTCCACCAAT 4039

Db 6597 GTTTCGCCACTGCTCTCTTCAAAAGCTCGACTTCACCTCAAGGAGGAGTAAATCGCTG 6656

QY 4040 GTTTCGCCACTGCGAGCCCGCAGAGGCGCGCTGCACTGCGAGGCGCGAGCAACGCTG 4099

Db 6657 GAGACCTCAGGTGAATTAATCCAAAGAGTGGCTGCAAGTGAGACTTCCAGAAAGCAATGAA 6716

QY 4100 CGGCCCCAGGTGAACACCCCAAGGAGTGGCTGCAGGTGGACTTCCAGAAACCATGAA 4159

Db 6717 AGTCACAGGAGTAATCTACTCAGGAGTAAATATCTCTGTTACCGAGCATGTATGTGAAGA 6776

QY 4160 GTGACTGGGCTGACCAACCGCGGCTCAAGAGCGCTGCTGACCGCATGTACGTGAAGA 4219

Db 6777 GTTCTCTCTCTCCAGCAGTCAAGATGGCCATCAGTGGACTCTCTTTTTTTCAGAAATGCCAA 6836

QY 4220 GTTCTCTGATCAGCAGCAGCGGAGCGGCAACCAAGTGGAGCCCTGTTCTTTCCAAAACGGCAA 4279

Db 6837 AGTAAAGCTTTTTCAGGGAATCAAGACTCCTTGCACCTGTGGTGAATCTCTTAGACCC 6896

Qy	1464	CTGCTGATCATCTTCAAGACACGAGCCAGCAGGCGCTTACACATCTACCCCGGATC	1523
Db	1498	ACTGATGTCGCTCTTGTATTCAAGACATTAACAAAGGTGTAAACATTTGAAGAT	1557
Qy	1524	ACCGAGTGGCGCCCTCTACAGCGCGCGCTGCCAAGGGCGTGAAGACCTGAGGAC	1583
Db	1558	TTTCCAAATTCGACGAGAAATTAATAATTAATGACAGTACTGTAGACATCGG	1617
Qy	1584	TTCCCATCTCGCGCGGAGATCTCAAGTACAGTGAAGCGGTGACCGTGGAGACGGC	1643
Db	1618	CCAACTAAATCAGATCTCGCTGCTGACCGCTATTACTCTAGTTTCGTTAATATGAG	1677
Qy	1644	CCCAACAAAGACGACCGCGCTGCTGACCGGCTACTACAGACGCTTCGTGAACATGAG	1703
Db	1678	AGAGATCTAGCTTACGAGACTCATTTGGCCCTCTCTCATCTGCTACAAAGATCTGTAGAT	1737
Qy	1704	CGCAGCTGGCCTCGGACTGATCGGCGCGCTGCTGATCTGCTACAGGAGAGCGTGGAC	1763
Db	1738	CAAGAGAAACACATATCTCAGACAGAGGAATGTATCTGTTTCTGTATTATTTGAT	1797
Qy	1764	CAGCGCGCAACCATGATCATGAGCAGCAAGCGCAAGCTGATCTGTTTACGGCTGTTCGAC	1823
Db	1798	GAGAACGAAAGCTGTGCTCCTCAGACAGAAATATACACGCTTTTCCCAATCCAGCTGA	1857
Qy	1824	GAGACCGCAGCTGTATCTCAGCAGACATCCAGCGCTTCTGCGCCACCGCTGGC	1883
Db	1858	GTGAGCTTGAGGATCAGAGTCCAGAGCTCCACATCATGACAGCATCAATGGCTAT	1917
Qy	1884	GTGAGCTGGAAGATCCGAGTTCAGGCGCAGCAACATCATGACAGCATCAAGCGCTAC	1943
Db	1918	GTTTTGTATGATTTGCAGTCTCAGTTTCTTTGATCAGGTGGCATCTGTACATCTTA	1977
Qy	1944	GTGTTCCAGACGCTCGAGCTGAGCGTGTGCTGATGAGGTGGCTACTGTACATCTCTG	2003
Db	1978	AGCATGAGACAGACATGACTTCTGCTTCTGCTTCTCTGATGATATACCTTCAACAC	2037
Qy	2004	AGCATGCGCGCCAGACGACTTCTCTGAGCGTGTCTTCTCGGGGTATACCTTCAAGCAC	2063
Db	2038	AAATGCTGTATAGACACATCACCTTATCCCATTTCTCAGGAGAACTGTCTTCATG	2097
Qy	2064	AAGATGCTGTACAGACACCTGACCTGTCTCCCTTCTCGCGGAGACTGTCTCATG	2123
Db	2098	TGATGAGAAACCCAGCTTATGATTTCTGGGTGGCCACAACTCAGACTTTCGGAACAGA	2157
Qy	2124	TCTATGAGAACCCCGGCTGTGATCTGAGTCTGAGTCTGACAAAGACGACTTCCGCAACCGC	2183
Db	2158	GGCATGACCGCCTTACTAGGTTCTAGTCTGACAAAGACACTGCTGATTATTACGAG	2217
Qy	2184	GGCATGACCGCCTGCTGAAGTCTCAGCTGCGAAGACACCGCGGCTACTACTAGGAC	2243
Db	2218	GACAGTTATGAAGATTTAGCATACTTCTGCTGAGTAAACAAATGCCATTGACCAAGA	2277
Qy	2244	GACAGTACGAGGACATCTCGGCTACCTGCTGTCCAAAGACAGCGCATCGACCGCGC	2303
Db	2278	AGCTTCTCCAGAAATTCAGAACACCTTACGACTAGGCAAAAGCAATTAATGCCAC	2334
Qy	2304	TCTTCTCCCAAACTCCCGCGCACCCAGCAGCGGTGACAGCAGTTCAACGCCAC	2360

RESULT 5
 LOCUS A05328 8241 bp RNA PAT 03-MAY-1993
 DEFINITION Synthetic Factor VIII cDNA insert of pCLB89.
 ACCESSION A05328
 NID 9345061

KEYWORDS	unidentified.
SOURCE	unidentified
ORGANISM	unclassified.
REFERENCE	1 (bases 1 to 8241)
AUTHORS	
JOURNAL	
FEATURES	Patent: WO 8809813-A 7 15-DEC-1988;
source	Location/Qualifiers
	1. .8241
	/organism="Artificial sequences"
	7. .7062
CDS	/codon_start=1
	/product="Factor VIII"
	/db_xref="Factor VIII"
	/translation="MQIELSTCFCLILRCEFSATRYILGAVELSHWDYMQSDLGELP
	VDARFPRVPKSPFFNTSVVYKKTLEVEFDHLFNIARQPPMGLGPTIQEYDVT
	VVITLKNMASHPVSLHAGVSYKASGEAYDDQTSQEKEDDKVFGGSHTYVQVL
	KENCPMADPLCLTYSLVSHVDLVKDLNSGLIGALLVREGSLAKEKTTQLHKEFLLF
	AVDEGRSHSETKNSLMQDRDAASARAMPKMHVYNGVYVNRSLGLGICHRKSYTWHV
	ICMGTTPEVHSIFLEGHTELVNRHQASLEISPIITFLTAQTLMLDGLQFLFCHSSH
	IQHDCMEA YKVDSCPEEPOLRMKNNEAEYDDDLTDSEMDVVRFDNDNSPFIQIRS
	VAKGHPKTVWHYIAAEEEDWDYAPLVAPDDRYSKSYLNGPQRIGKTKYKRFMAY
	TDFTFKTREAQHEESGILGLPYGEVGDVTLIIIFKNOASRPNIYPHGITDVRLYSR
	ILPKGVKHLKDFIPLCEIFKYVTVTVEGPTKSDPRCLTRYYSFVWNERDLASGL
	IGPLLIICYESVDQRQNIHSDKNVILFVFDENRSHWLTENIQRFILPAGVQIED
	PEQASNIMHSINGYVDSLSQVSLCHEVAYWYILSIGAQDTFLSVFFSYTFKRWV
	SYEDTILFPPFGSETVFMENPGLWILGCHNSDFRNRGMTALIKVSSCDKNTGDTYED
	PKQNVSSDDLMLILLRSPTHGLSHPDSTROKQFNATTIPENDIEKTDWFAHRTPM
	RQLHHSQDMVFTPEGLQLRLNEKLGTAATELKLKDFKVSSTNNLISITIPSONLA
	AGDNTSILGPPSNPVYDSQDITTLCKKSSPLTESGCPLSLSEENDSKLLEGLM
	NSQESSGNKVSSTESGRLFKGRAGHPALLTKDNALFKVSLSLKTKNTSNNSATNR
	KTHIDGFSILLIENSPPWQNIILESDTEFKKVTPLIHDRLMDKNATALRLNHSNKT
	SSKNMVMQKKEGPIPPDAONPDMGFKMLFLPESARW IORTHGKNSLNSGQSPSK
	QLVSLGPEKSEVQNFLEKNKVVVGKGFTHQVGLKEMVFPSSRLFLNLDNLHEN
	NTNHOKEITOEIEKKTILQENYVLPQIHTVGTGTFMKNLFLLSLTRQNVGEGSYDA
	YAPVLQDFRSLNDSNTKTKHTAHFKKGEENLEGIGNOTKQIIVEKYACTTRISPT
	SQNFVQSRKALKRFLPLEETELEKRIIVDDTSTQMSKNKKLITPSTLITQIDYNE
	KEKGATISPLSDCSLTHSPHOANRSPPLIAKVSFPSTRIPYIYLVFLQDNSSHL
	AASVKRKSQVQESLHSLQAKNNLSLAILTLEMTGDQREVGSIGTSATNVTYKRV
	ENTVLPKPLPKTSKVELLPKVIYQKDLFTETSNKSGPHGLDIVEGSLAQGTGAI
	KWENRNPGRVFFLRVATESSAKTPSKLLDPLAMDNHYGTQIPKEEKMSQKSPKTA
	FKKDDTILSLNACSNHIAAINEGONKPEIEVTWAKGRTERLCSONPPVLKRHORE
	ITRTTLQDQEEIDYDDTISVEMKEDFDIYDEENQSPRSFOKTRHYFIAVERLW
	DYGMSSSHVLRNRAQSGSVFQKVVVFQETDGSFTQPLYRGEINELHGLLGPVIRA
	EVEDNTMFTNRQASRPYSFYSLSIYEEDQROGAEPKRFKVPKNTKTFKRVQHM
	APTQDEEDCKAWAYESDVLDKDVHSLGLIPLLVCHTNTLNPAHGRQVTVQEEALFT
	IFDETSTWFTENNERNCRAPNTQMEDPTFKENTRPHAINGTIMDITPLGLVADQQR
	IRVYLLSGNSNETIHSHESGHFTVRRKEKYMALNLYPGVFETVEMLPKAGIR
	VECLIGELHAGNSTLIVYNSKQTPICMAWSGHIRDFOITASGOYQWAPKILARLY
	SGSINAWSTKPSFWIKVDLAPMIHIGINTQAGARQKFSLSIISQFIIMSLDGGKQW
	TYRGNSTFLMWFFGNVDSGSGKNIINFPPIIARYIRLPHTHYSIRSTLRMKNMGCDL
	NCSMP LGMESKAI SDAQITASSYVLTNMFATWSPSKARLHLOGRSNAVRPQVNPKEK
	LQVDFQKTHKVTGTTGGVKSFLTSYVYKFEFLISSQDGHQWTLFFQNGKVKVFGQNG
	DSFTPVNSLDPLLTLYRLIRHPQSVMHQIALRMEVLGCEADQDLY"

BASE COUNT 2625 a 1759 c 1703 g 2154 t
 ORIGIN
 Query Match 24.4%; Score 1085; DB 18; Length 8241;
 Best Local Similarity 73.8%; Pred. No. 0.00e+00;
 Matches 1681; Conservative 0; Mismatches 596; Indels 0; Gaps 0;
 Db 64 GCCACGAGAGATACCTACCTGGTGGCTGCACTGCACTGTCATGGCGACTATATCCAACTGAT 123
 ||||| | | ||||| ||||| | ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db	604	GA	GGGAGTCTGGCCACGAAAGAACA	CAGACCTTGCACAAATTTAT	TACTACTCTTTTGC	663
Qy	624	GAGGCGAGCTGGCCAGGAGAAA	CCGACAGCTTGCACAAGTTCA	TCCTGCTGTTTCGGC	683	
Db	664	GTATTTGATGAAGGAAAAGTTGG	CACTCAGAAACAAGAACTCCTT	GATGCGAGATAGG	723	
Qy	684	GTGTTTCAGCGGGGAAAGAGCT	GGCACACGAGACTAAGACAGC	CTGATGCGAGGACGC	743	
Db	724	GATGCTGCATCTCTCGGGCCTGG	CCCTAAATGACACAGCTCAAT	TGGTTATGTAAACAGG	783	
Qy	744	GACGCCACAGCGCCGCGCTGG	CCCAAGATGACACCGTTAAC	GGGTACGTGTAACCGC	803	
Db	784	TCTCTGCGAGGCTCTGATTGG	ATGCACAGAAATCACTCTATT	TGGCATGTGATTGAATG	843	
Qy	804	AGCTGCGCGGGCTCATCGGTG	CGACGGAAGCGGTGATGCGC	AGTCATCGGCATCGGCATG	863	
Db	844	GGCACCACCTCTGAAGTGCACT	CAATATTCTCGAAGGTACACA	ATTTCTTTGTGAGAAC	903	
Qy	864	GGCACCACCCCTGAGGTGCAC	AGCATCTTCTTGAAGGGCCAC	CTTCTCGTGCGCAAC	923	
Db	904	CATGCCACGGCTCTTGGAAATCT	CGGCAATACTTCTTACTTGCT	TACTGCTCAACACTTTG	963	
Qy	924	CACGGCAGCCAGCGCTGGAG	CATCGCCCATCACTTCTGACT	TTCGCCAGACCTGCTG	983	
Db	964	ATGACCTTGGACATTTCTACT	TGTTTTGTCAATCTCTTCCC	ACCAACATGATGGCATG	1023	
Qy	984	ATGGACCTAGGCCAGTTCTCT	GCTTCTGCCACATCAGGAC	CCACACGACGCGCATG	1043	
Db	1024	GAGCTTATGTCAAAGTAGACGT	CTCCAGGAGACCCCACTACG	AANTCAAAATAAT	1083	
Qy	1044	GAGCTTTAGTGAAAGTGACAG	CTGCCCGGAGCGCGCTCGC	GATGGAACAAC	1103	
Db	1084	GAAAGACGGCAACATATGAT	GATCTTACTGTAATGATGGT	TGTCAGGTTT	1143	
Qy	1104	GAGAGGCGAGGATACGACG	ACACCTTGACGACGATGGAT	GTGTCAGCTTC	1163	
Db	1144	GATGATCAGAACTCTCTCT	TTATTCAAATTCGCTAGTT	CGCCAAAGCAAGCATCTAA	1203	
Qy	1164	GACGAGCAACAGCCCGACGT	CTATCCAGATCGCGAGCTGG	CCAAAGACACCTAAG	1223	
Db	1204	ACTTGGGTACATTATGCTG	TGAAGAGGAGCTGGGCA	TATGCTCCCTTAGTCTCT	1263	
Qy	1224	ACCTGGGTGCATACATCG	CGCCGAGGAGGAGCTTGG	GACTACGCCCGCTAGTACTG	1283	
Db	1264	CCCCCGATGACAGAGTTATAA	AGTCAATTTGAACATGGCC	CTCAGCGGATTTGGT	1323	
Qy	1284	CCCCCGACGACCGGACG	ACTACAGAGCCAGTACCT	GAAACGGCCCCGACGCA	1343	
Db	1324	AGGAAGTACAAAAAGTCCG	AGATTTATGGCATACACAG	ATGAACCTTTAAGACTCTG	1383	
Qy	1344	CGAAGTACAGAAAGTGCG	CTTATGGCCTACACCG	ACGAGCTTTCAAGACCCCG	1403	
Db	1384	GCTATTACGATCAATCAGGA	ATCTTTGGACCTTTACTTT	TATGCGGAAGTTGGACACA	1443	
Qy	1404	GCATCCAGCAGATCTCGG	GCATCTCGGCCCTTGTAC	CGGAGGTGGCGACAC	1463	
Db	1444	CTGTTGATTATATTAAGAA	TCAAGCAAGCAGACGACAT	ATAATCTACCTCAGCGA	1503	
Qy	1464	CTGCTGATCATCTTCAAG	ACCAACGAGCCGACGAG	CCCTACACATCTACCC	1523	
Db	1504	ACTGATGTCGCTCTTCTAT	TATTAAGAGATTTACAA	AGGTGTAAACATTTGAG	1563	
Qy	1524	ACCGAGTCGCGCCCTCTG	ACGCGCGCCCTGCCCA	AGGGCTCAAGACACTTGA	1583	

[illegible]

RESULT	7
LOCUS	E00527 8831 bp RNA PAT
DEFINITION	Complete cDNA encoding human factor VIII.
ACCESSION	E00527
NID	g2168806
KEYWORDS	JP 1985243023-A/2.
SOURCE	Homo sapiens.
ORGANISM	Homo sapiens
	Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
	Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae

Qy	2244	GACAGCTAGGAGGAGCATCTCGGCTTACTGCTGTCTCCAGACAGCAACAGCCGATTCGAGCCCGCC	2303
Db	2387	AGCTTCTCCGAGAAATTCAGACAGACCCCTAGCACTAGGCAAAAGCAATTTTAATGCGCAC	2443
Qy	2304	TCCTTCTCCAAAATCTCCGCGCCAGCCCGACGCGCTCAGAAAGAGTTCAACGCCAC	2360
RESULT	8		
LOCUS	HUMFVIII	9029 bp	mRNA
DEFINITION	Human coagulation factor VIII:C mRNA, complete cds.		
ACCESSION	M14113		
NID	g182817		
KEYWORDS	coagulation factor VIII:C.		
SOURCE	Human kidney, cDNA to mRNA, clones pF8[100,102]; and cell line GM1416 DNA, clone pF8-4.3.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryotes; mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Euthieria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	1 (bases 1 to 9029)		
	Truett,M.A., Blacher,R.W., Burke,R.L., Caput,D., Chu,C., Dina,D., Hartog,K., Kuo,C.H., Maslarz,F.R., Merryweather,J.P., Najarian, Pachtl,C., Potter,S.J., Puma,J., Quilroga,M., Rall,L.B., Randolph Urdea,M.S., Valenzuela,P., Dahl,H.-H.M., Favalaro,J., Hansen,J., Nordfang,O. and Ezban,M.		
TITLE	Characterization of the polypeptide composition of human factor VIII:C and the nucleotide sequence and expression of the human kidney cDNA		
JOURNAL	DNA 4 (5), 333-349 (1985)		
MEDLINE	86081164		
REFERENCE	2 (sites)		
AUTHORS	Youssefian,H., Wong,C., Aronis,S., Platokoukis,H., Kazazian,H. Jr. and Antonarakis,S.E.		
TITLE	Moderately severe hemophilia A resulting from Glu----Gly substitution in exon 7 of the factor VIII gene		
JOURNAL	Am. J. Hum. Genet. 42 (6), 867-871 (1988)		
MEDLINE	88220354		
REFERENCE	3 (sites)		
AUTHORS	Bernardi,F., Legnani,C., Volinia,S., Patracchini,P., Rodorigo,G. DeRosa,V. and Marchetti,G.		
TITLE	A HindIII RFLP and a gene lesion in the coagulation factor VIII gene		
JOURNAL	Hum. Genet. 78 (4), 359-362 (1988)		
MEDLINE	88197150		
COMMENT	[3] sites; mutation causing hemophilia. [2] sites; mutations causing hemophilia. Draft entry and clean copy sequence for [1] kindly provided by M.A.Fruett, 26-FEB-1986. The mutation at position 1042 results in a change of Glu to Lys and the one at position 1043 in a change from Glu to Gly. The mutation at position 1042 produces a premature stop codon.		
FEATURES	Location/Qualifiers		
source	1. .9029 /organism="Homo sapiens" /db_xref="taxon:9606" /map="Xg28"		
mRNA	1. .9029 /note="F-VIII mRNA"		
gene	172. .7227 /gene="F8C"		
sig_peptide	172. .228 /gene="F8C"		
CDS	/note="Coagulation factor VIII:C signal peptide" 172. .7227 /gene="F8C"		

Qy	2244	GACAGCTAGGAGGAGCATCTCGGCTTACTGCTGTCTCCAGACAGACAAAGCCATTTGAGCCCGCCG	2303
Db	2387	AGCTTCTCCGAGAAATTCAAGACAGCCCTAGCACTAGGCAAAAGCAATTTTAATGCGCACG	2443
Qy	2304	TCCTTCTCCAAAACATCCCGCCAGCCCGACGCGCTCAGAAAGAGTTCAACGCCACG	2360
RESULT	8		
LOCUS	HUMFVIII	9029 bp	mRNA
DEFINITION	Human coagulation factor VIII:C mRNA, complete cds.		PRI
ACCESSION	M14113		
NID	g182817		
KEYWORDS	coagulation factor VIII:C.		
SOURCE	Human kidney, cDNA to mRNA, clones pF8[100,102]; and cell line GM1416 DNA, clone pF8-4.3.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryotes; mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Euthieria; Primates; Catarhyni; Hominidae; Homo.		
AUTHORS	1 (bases 1 to 9029)		
	Truett,M.A., Blacher,R.W., Burke,R.L., Caput,D., Chu,C., Dina,D., Hartog,K., Kuo,C.H., Maslarz,F.R., Merryweather,J.P., Najarian, Pachtl,C., Potter,S.J., Puma,J., Quilroga,M., Rall,L.B., Randolph Urdea,M.S., Valenzuela,P., Dahl,H.-H.M., Favalaro,J., Hansen,J., Nordfang,O. and Ezban,M.		
TITLE	Characterization of the polypeptide composition of human factor VIII:C and the nucleotide sequence and expression of the human kidney cDNA		
JOURNAL	DNA 4 (5), 333-349 (1985)		
MEDLINE	86081164		
REFERENCE	2 (sites)		
AUTHORS	Youssefian,H., Wong,C., Aronis,S., Platokoukis,H., Kazazian,H. Jr. and Antonarakis,S.E.		
TITLE	Moderately severe hemophilia A resulting from Glu----Gly substitution in exon 7 of the factor VIII gene		
JOURNAL	Am. J. Hum. Genet. 42 (6), 867-871 (1988)		
MEDLINE	88220354		
REFERENCE	3 (sites)		
AUTHORS	Bernardi,F., Legnani,C., Volinia,S., Patracchini,P., Rodorigo,G. DeRosa,V. and Marchetti,G.		
TITLE	A HindIII RFLP and a gene lesion in the coagulation factor VIII gene		
JOURNAL	Hum. Genet. 78 (4), 359-362 (1988)		
MEDLINE	88197150		
COMMENT	[3] sites; mutation causing hemophilia. [2] sites; mutations causing hemophilia. Draft entry and clean copy sequence for [1] kindly provided by M.A.Fruett, 26-FEB-1986. The mutation at position 1042 results in a change of Glu to Lys and the one at position 1043 in a change from Glu to Gly. The mutation at position 1042 produces a premature stop codon.		
FEATURES	Location/Qualifiers		
source	1. .9029 /organism="Homo sapiens" /db_xref="taxon:9606" /map="Xg28"		
mRNA	1. .9029 /note="F-VIII mRNA"		
gene	172. .7227 /gene="F8C"		
sig_peptide	172. .228 /gene="F8C"		
CDS	/note="Coagulation factor VIII:C signal peptide" 172. .7227 /gene="F8C"		


```
/note="preprocoagulation factor VIII:C"
/codon_start=1
/db_xref="GDB:600-119-124"
/db_xref="PID:g182818"
/translacion="MOELSTCFCLLLRFSFSTRYYIGAVELSDWYMSDGLGELP
VDAREPVRPKSPFNTSVYKYLTFVEFDHLENIAKRPWPWMLLGLPTQAEVDT
VITLRNMAHPVSLHVGYSYKMAEGAEYDQDTSQREKEDKVPFGSGSYTYWQVL
KNGPMSADPLCLITYSYLHVLDVLDLNSGLIIGALLVCREGLAKEKTOTLHKILIF
AVFDEKSHWSETKNSLMQDRDAASARAPKMTVNGYVNRSLPGLIGCHRSYVHV
ICMGTFEVHSIFLEGTFIVRNHROASLEISPIITFLTAQTLMDLQGLFLFCHSSH
QHDGNEAVVQDSQEPQLRMKNQAEEDYDDLDLTDSEMDWRFDDNSFSFIOIRS
VAKKHPKTWHYIAAEEDWDYAPIVLAPDDBSYKQYLNNQPORICRKYKVPFAY
TDETKREAIQHESGILGLPILGYVEGDTLLIFKNQASRPYNIYPHGITDVRLYSR
RLPKGVKHLKDPFLLPGEIFKRYKTVTVDEGPTKSDPRLTRYYSFVNEMERDLASGL
ICPLLCYKESVDQRGNQIMSKRNVILFSEVDENRSHYLTENIQRELPNPAQVLED
PEFAQSNIHISINGVFDISQVLSVCLHEVAYWYIILSIAQTDFLSVFFSGYTFKRWV
YEDTLTFPFSGETVFMENPGLITLGHNSDFRNRGMTALLKVSCKDNTGYED
SYEDISAYLLSKNNAIEPRPSFONSHRSTROKFNATTTIPENDIEKTDPFARHPM
PKIQNVSSDMLMLLRQSPPTPHGLISLQAEKYTEFSDPSGAIDSNNSLSMWHF
RPQLHSGDMVFTPEGSLQRLNKLGTAAATELKKLDPKFVSSNNLISTIPDNLA
ACTDNTSSLGPSPWPHVDLSQDITLFGKSSPLTESGGLSLSEENNDKILIESGLM
NQSESNGKRVNVSSEGLKFKRAHGPALLTDNALFKVSLILTKNTKSNNSATNR
KTHIDGSLIENSFSVMONTLESDFEKVTPLIHDMRLMDKNATALRLNHSNKT
SSKNEMVQKQEGP IPDAQNDMSFTKMLFLPESARWIORTHGKNSLNSQPSPK
QLVSLGPKSVQEQNLSEKRVVVGKGEFTKOVGLKENVFPSSRNLFILNLDNLHEN
NTHNQEKILQEEIEXKETLTOENWLPQIHTVTGTFKNFMNLFLLSTRQNVESYDGA
YAPVLQDFRSINDSTRNRYKHTAHFSKGEENLEGLGNQTKQIVKRYACTRISNT
SOQNVTVQSRKALQFRLPLEETELEKRIIVDDTSTQWSKNMKHLTPSTLQIDYNE
KKGAIQTSPSLDCLTRSHSIPQANRSLPIAKVSSFPISIRPYLTRVLFDQNSHLIP
AASVKKDSGVQSSHSFLQAKRNNLSAILIEMTCDQREVGSLGTSATNSVTYKV
ENTVLPKDLPTKSKVELLPKVHIYQKDLFPTETSNQSPGHLDLVEGSLIQCTGAI
KNEANRPKVPFLVAVESAKTSPKLLDPLAWDNHGTQIPKEEWSQERSPEKTA
FKKDTILISLACENSDIAAINEQONKEIEVTAQQRTERLCSQNPVILKRBHORE
ITRTTLASDQEEIYDDTIVSEMKKEDFDIYDEENQSPRSEFKTRHYETAAVERLW
DYGMSSPHVLRNRAQSGSPQPKVVFQETDGSFTQPLYRGEIHELGLLGPYIRA
EVEDINIVTRNQAARSPYSYSSLSYEEDQORGAEPKRNFKVNEKTYFWKYOHM
APTKDEDFKAMAYSDVDLEKDVHSLGILPVLVCHTNTLPNAGROVTVQEFALFT
IFDETKSWYTEMERNCRAPCNIQMEDPTKENYRFAINGYIMDTLPLGLVMAQDOR
IMYLLGSHGNENIHSHFSGHFTVRKKEKYMALYNDLPVETVEMLPE SKAGIWR
VECLIGELHAGKSTLFLVSNKQSTPLGMAASHIRDFOITASGOYQOMAPKARLHY
SGSINAMSTKEPSPKIVDLPAPMIIHGKIQGAKQKSSLYISQFIIMYSIDCKWQ
TYRGNSTGLTLMVFFGNVDSSGIKHNIENPPIIARYIRLHPHYISIRSLRMLMGDL
NSCSNPLGMESKALISDAQITASSYFTNMFATWSPSKARLHLQGRSNARPOVNNPKEM
LOVDFOKMTKVTGTTQGVKSLSTSNVYKEFLISSSDGQHWTLFTQNGKVKVQNGQ
DSFTPVVNSLDPPLLTRYLRIPHQSWVHQIALRMEVLGCEAQDLY"
898..7224
/ene="F8C"
/note="coagulation factor VIII:C"
1042
/ene="F8C"
/note="g in wt; a in hemophilic [2]"
1043
/ene="F8C"
/note="a in wt; g in hemophilic [2]"
6853
/ene="F8C"
/note="c in wt; t in hemophilic [3]"
BASE COUNT 2860 a 1910 c 1848 g 2411 t
ORIGIN 185 bp upstream of SacI site.
```

Query Match 24.4%; Score 1085; DB 21; Length 9029;
Best Local Similarity 73.8%; Pred. No. 0.00e+00;
Matches 1681; Conservative 0; Mismatches 596; Indels 0; Gaps 0;

```
Db 229 GCCACCAAGAGATCTACTGGTGCAGTGAACCTCATGGCACTATATGCAACTGAT 288
||||| | | ||||| ||||| | ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 84 GCCACCGCGCGTACTACTGGGCGCGGTGAGGCTGCTCTGGGACATACATGACAGGAC 143
289 CTCGGTGAAGCTGCTGGAAGCAAGATTTCTCTCTAGTGCCTCAAAATCTTTTCCATTC 348
||| ||||| | | ||||| | | ||||| | | ||||| | | ||||| | | ||||| |||||
Qy 144 CTGGGCGAGCTCCGCTGGAGCGCGCTTCCGCCCGCGGTGCCCAAGAGCTTCCCTTC 203
349 ACACCTCAGCTGCTGACAAAGACTCTGTTGTAGAAATACCGATCACCCTTTTCAAC 408
||||| | | ||||| | | ||||| | | ||||| | | ||||| | | ||||| |||||
Qy 204 AACACGAGCTGCTGTAAGAAAACCTGCTCGTGGAGTTTACCCACCACCTGTTCAAC 263
409 ATCGCTAAGCAAGCCCTGGATGGGTCTGCTAGGCTTCCATCCTACCATCCAGGTGAGTT 468
||| ||||| | | ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 264 ATTGCAAGCGCGCGCGCTGGAATGGGCTCTGGGCCCCACCATCCAGGCGGAGGTG 323
469 TATGATACAGTGGTCATTACACCTAAAGAACTGGCTTCCCATCTCTGCTCAGTCTTTCATGCT 528
||| ||| ||||| | | ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 324 TAGACACCGTGGTCATCACCTTGAAGAACATGGCCACCCCGTCAGCTTGCACGCC 383
529 GTTGTCTATCTCTGGAAGCTTCTCAGGGAGCTGAATATGATGATCAGACCACTCAA 588
||| ||| ||||| | | ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 384 GTGGCGCTGAGCTACTGGAAGGCCAGCGAGGCGCGGAGTACGACGACCACTGCCAG 443
589 AGGAGAAAGATGATGATGCTTCCCTGCTGGAAGCCATACATATGCTGGAGGTC 648
||||| | | ||||| | | ||||| | | ||||| | | ||||| | | ||||| |||||
Qy 504 CTTAAGGAGAGCGGCTATGCGCAGGACCCCTGCTGCTGACCTACAGTACCTGACC 563
709 CATGTGACCTGCTGTAAGACTTGAATTCAGGCGCTCATTTGAGCGCTTACTAGTATGAGA 768
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 564 CAGTGGACCTGGTGAGGATCTGAAACGGGGCTGATCGGGCGGCTGCTGGTGTGCGC 623
769 GAAGGGAGTCTGGCCAAAGGAAACACAGACCTTGCACAAATTTATATCTACTTTTGTCT 828
||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 624 GAGGCGAGCTGGCCAAAGGAAACCCACACCTGCAAGTTTCATCTGCTGCTTCCGCG 683
829 GTATTTGATGAGGAAAGCTGGGCTGCAAGAAACAAAGAACTCTTTCATGCGAGGATAGG 888
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 684 GTGTTGACGAGGGAAGAGCTGGCAGCGGAGACTAAGAAACAGCTGATGCGACCGCG 743
889 GATGCTGATCTGCTGGGCGCTGGCTTAAATGCAACAGTCAATGTTATGTAACAGG 948
||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 744 GACGCGCGAGCGCGCGCTGGCCCAAGATGACACCGCTTAACGGCTGCTGAACCGC 803
949 TCTCTGCCAGCTCATTTGATGATCCACAGGAAATCAGTCTATTGGCATGTGATGGAATG 1008
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 804 ACCTGCGCGGCTGATGGCTGGCCCGGAGAGCGGTGCTGAGGACGTCTCGGCATG 863
1009 GGCACCACTCTGGAAGTGCATCAATATTCCTGGAAGGTCACACATTTCTTGTGAGGAAC 1068
||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 864 GGCACCACTCTGAGGTGACAGACTTCTCTGGAGGGCCACACCTTCTCTGGTGGCAAC 923
1069 CATGCCAGGCGCTCTTGGAAATCTCGGCAATACTTTCCTTACTGCTCAACACTCTTG 1128
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 924 CACCGCCAGCGCAAGCTGGAGATCAGGCCCATCACCCTTCTGACTGCCGACAGCCCTGCTG 983
1129 ATGACCTGGAGAGTTTCTACTGTTTTTGTCTATATCTTTCCACCAACATGATGCGATG 1188
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 984 ATGACCTAGGCCAGTTCTCTGTTGTTGTCACATCAGCAGGCCACCGACGACGCGCATG 1043
```


Db	298	TATGATACAGTGGTCATTACACTTAAAGAACATGGCTTCCCATCTCTGTGCAGCTCTCATGCT	357
Qy	324	TAAGCAACCGTGGTGTATCACCTTGAAGAACAATGCCAGCCACCCGTCAGCGCTGCAGGCC	383
Db	338	GTTCGGTGTATCCTACTCGAAGAGTTCCTGAGGAGCTGAATATCATCATCAGCAACAGCTCAA	417
Qy	384	GTGGGCGTGAAGTCTGGAGGCGCCAGGAGGCGCCGAGTACGACGACGAGAGCTCCAG	443
Db	418	AGGCAAGAAAGATGATAAAGTCTTCCTGTCGAAGCCATACATATGTCTCGCAGGTC	477
Qy	444	CGCAGAGAGGAGCAGCAGAGTGTTCGCGGGGGAGCCACACCTACGTGTGGCAGGTG	503
Db	478	CTGAAGACAGAAATGGTCCAAATGGGCTCTGAGCCACTGTGCGCTTACCTACTCATATCTTTCT	537
Qy	504	CTTAAGAGAGCGGCCCTATGGCCAGCAGCCCTGTGCGCTGACCTACAGCTACCTGACG	563
Db	538	CATCTGACCTCGTAAAGACTTGAATTACGGCCTCATTTGGAGCCCTACTAGTATGTAGA	597
Qy	564	CACGTGCACTGGTCAAGGATCTGAACACGGGGCTGATCGGGGCCCTGCTGGTGTGCGC	623
Db	598	GAAAGGAGTCTGGCCAGGAAAGACACAGACCTTGCACAAATTTATACTACTTTTTGCT	657
Qy	624	GAGGCAAGCTGGCCAAAGAGAAACCCAGACCCCTGCACAACTTCATCTGCTGTTCCGC	683
Db	658	GTATTTGATGAAGGAAAGTGTGGCACTCAGAAACAAAGACCTCCTTCGATCGAGATAGG	717
Qy	684	GTGTTGACAGGGGACAGCTGGCAGCGAGACTAAGAACAGCCTGATGCGAGGACCG	743
Db	718	GATGCTGCATCTGCTGGGCTGGCCATAAATGCACACAGCTCAATGGTTATGTAAACAGG	777
Qy	744	GACGCGCCAGCGCCGCGCTGGCCGAGATGCACAGCTTAACGGCTACGTGAACCGC	803
Db	778	TCTCTCCAGGCTCTGATTGGAATGCCACAGGAATACGCTCTATTGGCATGTGATTGGAATG	837
Qy	804	AGCTTGGCGGGCTGATGGCTGCCACGGAAGCGGTGATGCGCAGCTATCGGCATG	863
Db	838	GGCACCACTCTCGAAGTGCACATAATTTCTCGAAGGTGCACATACTTTCTGTGAGAAC	897
Qy	864	GGCACACCCCTGAGGTGCACAGCATCTCTCGAGGGCCACACCTTCCTGTCGGCGAAC	923
Db	898	CATCGCCAGCGCTCTTGGAAATCTCGCCCAATACTTTCTTACTGCTCAACACTCTTG	957
Qy	924	CACGGCCAGCCAGCTGGAGATCAGCCCACTACCTTCTGACTGCCAGACCTTGCTG	983
Db	958	ATGACCTTGGACATTTCTACTGTTTTGTCTATATCTTCTCCCAACATGATGGCATG	1017
Qy	984	ATGGACCTAGGCCACTCTCTGCTTCTGCCATCATCAGCAGCCACGACGAGCGGCATG	1043
Db	1018	GAGCTTTATGCAAGTAGACAGCTGTCCAGAGAACCCCACTACGAATGAATAAT	1077
Qy	1044	GAGCCTTACGTGAAGTGGACACTGCCCGGAGGACCCCACTGCGCATGAAGACACAC	1103
Db	1078	GAAAGCGGAAAGATCATGATCATCTTACTGATTTCTCAAAATGCATTTGGTCAGTTTT	1137
Qy	1104	GAGAGCGCGAGGACTAGCAGCAGACCTGACCGACAGAGATGGATGTCTGTCAGCTTC	1163
Db	1138	GATGATGACAACTCTCCTCTCTTATTCAAAATTCGCTTCAGTTGCCAAGACAGCTCTAAA	1197
Qy	1164	GACGACGACACGCGCCAGCTTCATCCAGATCGCAGCGTGGCCAGAGACACCTAAG	1223
Db	1198	ACTTGGGTACATTATGCTGCTGAAGGAGGACTGGGAGCTATGCTCCCTTAGTCTCTC	1257
Qy	1224	ACCTGGTGCATCATCTGCGCGCAGAGAGGACTGGGACTAGCGCCGCTAGTACTG	1283

Db	1258	GGCCCGGATGACAGAAAGTTATAAAAGTCAATATTTGAAACAATGGCCCTTCAGCGGATTTGGT	1317
Qy	1284	GGCCCGGAGACGGCAGCTACAAGACCCAGTACCTGAAACAACGGCCCCCAGCGCATCGGC	1343
Db	1318	AGGAAGTACAAAAGATCCCGATTATATGGCATAACAGACTCAAAACCTTTAAAGACTCTGAA	1377
Qy	1344	CGCAAGTACAGAAGGTGCGCTTCAATGCGCTACACGACGAGACTTTTCAAGACCCCGCAG	1403
Db	1378	GCTATTCCAGCATGAATCAGGAATCTTTGGGACCTTTACTTTATTTGGGGAAGTTGGACACACA	1437
Qy	1404	GGCATTCAGACACAGTTCGGGATCATCTCGGCCCCCTGCTGTACGGCGAGGTGGGCGACACC	1463
Db	1438	CTGTGTATTATATTTAAGAAATCAAGACGACAGACATATAACATCACTACCTCTCAGCGAATC	1497
Qy	1464	CTGCTGATCATCTTCAAGACACAGGCCGACAGGCCCTCAACATCTACCCCAACGGGATC	1523
Db	1498	ACTGATGTCCGTCCTTTGTATTCAAGGAGATTACAAAGAGTGTAAACAATTTGAAGAT	1557
Qy	1524	ACCGAGCTGGCGCCCTCTGACGCCCGCCCTGCCCAAGGGCGTCAAGACCACTTGAGGAC	1583
Db	1558	TTTCCAATTCTGCCAGGAGAAATTTCAAAATATAAATAGACAGTACTGTAGAGATGGG	1617
Qy	1584	TTCCCATCTCTCGCGCGGAGATCTTCAAGTACAAGTGGACCGTCACTGCGTGGAGACGCG	1643
Db	1618	CGAATCAATACATCTCTGGTGCCTGACCGGCTATTACTAGTTTTCGTTAAATATGAG	1677
Qy	1644	CCCACCAAGGACGCCCGCTGCTGACCGGCTACTACAGCAGCTTCTGTAACATGGAG	1703
Db	1678	AGAGATCTAGCTTCAGGACTATTTGGCCCTCTCTCATCTGCTCAAGAATCTGTAGAT	1737
Qy	1704	CGGACCTGGCCTCGGACTGATCGGCCCTCTGCTGATCTGCTACAGAGAGCGGTGCAC	1763
Db	1738	CAAGACGAAACCCAGATAATCTCAGACAGAGAGTAATCATCTCTGTTTCTGTATTGAT	1797
Qy	1764	CAGCGCGGCAACCATCATGACGGCAAGCGCAAGTGAATCCTGTTACACGGTGTTCAC	1823
Db	1798	GAGAACCGAGCTGGTACTCTCAGACAGATATACAAGCGTTCTTCCCAATCCAGCTGGA	1857
Qy	1824	GAGAACCGCAGCTGGTATCTGACCGAGAAACATCCAGCGTCTCTGCCCAACCCGCTGCG	1883
Db	1858	GTGAGCTTGAGGATCAGAGTTCCAAAGCCTCBAACATATGCACAGCATCAATGGCTAT	1917
Qy	1884	GTCCAGCTTGAACATCCGAGTTCCAGGCCAGCAACATCATGCACAGCATCAACGCGTAC	1943
Db	1918	GTTTTTGATAGTTTCGATGTTCAGTTGTTTGCATGAGGTGGCATCTGCTACATTTCTA	1977
Qy	1944	GTGTTTCGACAGCCTGCAGCTCAGCGTGTGCGTGCATCAGGTGGCCTACTGTACATCCTG	2003
Db	1978	AGCATTTGAGACACAGACTGACTCTCTTTCTGCTTTCTCTCTGATATACCTTCAACAAC	2037
Qy	2004	AGCATGGCGCCCAACACCGACTCTCTGAGCGTGTCTTCTCGGGGTACTACTTCAGCAC	2063
Db	2038	AAATTTGTTCTATCAAGACACATCAACCTTATTTCCCATTTCTCAGGAGAAATCTGTCTCATG	2097
Qy	2064	AGATGTGTATCAGACGACACCTTGACCTGTGTTCCCTTCTCCGGCGAGACTGTGTTCA TG	2123
Db	2098	TCGATGGAAAACCGAGTCTATGGAATCTCGGGTGGCAACAACCTCAGACTTTTCGGAACAGA	2157
Qy	2124	TCATATGAGAAACCCCGCCTGTGGAATCTCGGGTGGCAACAACGAGCACTTCGGAACCGC	2183
Db	2158	GGCATGACCCCTTACTGAAGCTTTTCTAGTTGTGTACACAAGCAACTGCTGATTATTACGAG	2217
Qy	2184	GGCATGACTGCCCTGCTGAAGTCTCCAGCTGCGCAAGAACCAACGCGGCACTCTACGAG	2243

Db 748 GAAGGAGCTGCGCAAGGAAAGACACAGACCTTCACAAATTAATACTACTTTTTGCT 807
Qy 624 GAGGCGAGCTGGCCAGAGGAAACCCAGACCCCTGCACAGTTCACTTCGTGTCGCC 683
Db 808 GTATTGATGAGGCAAAAGTTGGCACTCACAACAAGAACTCCTTGATGACGATAGG 867
Qy 684 GTGTTGACGAGGGGACAGCTGGCAGCGAGCACTTAACAACAGCCTGATGACGACCGC 743
Db 868 GATGCTGCATCTGCTCGGGCTGGCTGAAATGCACAGACTCAATGTTATGTAACAGG 927
Qy 744 GAGCGCGCAGCGCGCGCTGGCCCAAGATGCACACGCTTAACGGCTACGTGACCGCG 803
Db 928 TCTTGCCAGGCTCTGATTGGATGCCACAGGAAATCACTATTGGCATGTGATGGAATG 987
Qy 804 AGCGTGGCGGGCTGATCGGCTGCACCGCAAGGGGTGACTGGCAGCTCATCGGCATG 863
Db 988 GGCACCACTCCTGAAGTGCACCTCAATATTCGAAAGGTGCACACATTTCTTGTCAGGAAC 1047
Qy 864 GGCACCACTCCTGAGGTGCACAGCATCTTCTTGAGGGCCACACTTCTGCTGGTGGCGAAC 923
Db 1048 CATGCCAGCGCTCTTGGAAATCTGCCCAATCACTTTCTTACTGCTCAACACACTCTTG 1107
Qy 924 CACCGCCAGCGGCTGGAGTCAAGCCCATCACTTCTTCACTGCTGCCAGACCTCTG 983
Db 1108 ATGGACCTTGGACATTTCTACTGTTTGTGATATCTCTCCAGACCAATCATGCGCATG 1167
Qy 984 ATGGACCTAGGCGACTCTCTGCTTCTGTCACATCAGCGCCACACGACGACGCGCATG 1043
Db 1168 GAAGCTTATGCAAGTAGACAGCTGTCAGAGGAACCCCAACTCAAGATGAATAAATAT 1227
Qy 1044 GAGGCTTACGTGAAGTGGACAGCTGCCCCGAGGAGCCCGAGCTGCGCATGAGAAACAC 1103
Db 1228 GAAGAGCGGAAGCATGATGATGATCTTACTGATTTCTGAATGGAATGCTGTCAGGTTT 1287
Qy 1104 GAGGAGCGGAGGACTACGAGGAGCACTGACCGACGCGAGATGAGATGCTGACGGTTC 1163
Db 1288 GATGATGACAACTCTCCTTCTTTATCAAAATTCGCTCAGTTGGCAAGAGCATCTAAA 1347
Qy 1164 GACGACGACAAACGCCCGAGCTTCATCCAGATCCGACGCTGGCCAAAGACCCCTAAG 1223
Db 1348 ACTTGGGTACATTAATTTGCTGCTGAAGAGGAGCTGGGACTATGCTCCCTAGTCTCTC 1407
Qy 1224 ACCTGGGTGCACTACATCGCGCGCGAGGAGGAGCTGGGACTACGCCCCGTACTACTG 1283
Db 1408 GCCCGGATGACGAAGTTATAAAGTCAATATTTGAACAATGGCCCTCAGCGGATTTGGT 1467
Qy 1284 GCCCGGACGCGGCTACAGAGCCAGTACCTGAACAGCGGCCCGCCAGCGCATCGGC 1343
Db 1468 AGGAAGTCAAAAAGTCCGATTTATGGCATACACAGATGAACCTTTAAGACTCGTGAA 1527
Qy 1344 CGCAAGTACAAAGAGTGGCTTCATGGCTTACACCGACGAGACTTTCAAGACCGGGAG 1403
Db 1528 GCTATTGACATCAATCAGGAATCTTGGAGCTTTACTTTATGGGAGTTGGAGCACA 1587
Qy 1404 GCCATCCAGCAGCTGGCGGCTCTCGGCGCCCTGCTGTACGGCGAGTGGGCGACAC 1463
Db 1588 CTGTTGATTATTAAGATCAAGCAGGACGACCATATATACATCTACCTCAGCGAATC 1647
Qy 1464 CTGCTGATCATCTTCAAGACCGGCGCAGGAGCCCTACAACTATACCCCGCAGGCATC 1523
Db 1648 ACTGATGCTGCTCTTGTATTCAGGAGATTAACAAAAGGTGTAACAATTTGAAGAT 1707
Qy 1524 ACCGACGTGGCGCCCTGTACAGCGCGCCCTGCCCAAGGGGCTGAAGCACCTGAGGAC 1583

Db 1708 TTTCGAATCTGCCAGGAGAAATATCAAAATATAAATGACAGTCACTGTAGAAGATGG 1767
Qy 1584 TTCCCATCTCTGCCCGGAGATCTTCAAGTAAAGTGGACCGTGACCGTGGAGACGGC 1643
Db 1768 CCAACTAAATCAGATCTCGGTGCTGACCGCTATTACTCTAGTTCGTTTATATATGAG 1827
Qy 1644 CCCACCAAGAGCGACCCCGCTGCTGACCGCTACTACAGCAGCTTCTGTGAACATGGAG 1703
Db 1828 AGACATCTAGCTTCAGGACTCATTTGGCCCTCTCTCATCTGCTCAAGAAGATCTGTAG 1887
Qy 1704 CGCGACCTGGCTCCGACTGATCGGCCCTCTGATCTGCTACAAGGAGCGCTGGAC 1763
Db 1888 CAAAGAGGAAACCAATATGTACAGACAGAGAAATGTCACTCTTTCTGTATTTGAT 1947
Qy 1764 CAGCGGGCAACCAAGATCATAGGCAAGCAAGCGATGATCTGTCAGCGTGTTCGAC 1823
Db 1948 GAGAACCGAAGCTGTACTCACAGAGAAATATACAAGCTTTCTCCCAATCCAGCTGA 2007
Qy 1824 GAGAACCGAGCTGGTATCTGACCGAGACATCCAGCGCTTCTGCCCAACCCCGTGGC 1883
Db 2008 GTGCAGCTTGAGATCCAGAGTTCCAAAGCCTCCAACATCATGCAGCATCAATGGCTAT 2067
Qy 1884 GTGCAGCTGGAAGATCCCGAGTTCCAGGCGACCAACATCATGCAGCATCAACGGCTAC 1943
Db 2068 GTTTTTCATGTTTGCAGTTGCTCATGTTTGTTCGATCAGGTGGCATACTGGTACATTCTA 2127
Qy 1944 GTGTTTCAGACGCTGCAGCTGAGCGTGTGCTGATGAGGTGGCCTACTGCTACATCCTG 2003
Db 2128 AGCAATGGAGACAGACTGACTTCTCTTCTGCTTCTCTCTGATATACCTTCAACAC 2187
Qy 2004 AGCAATGGGCGGACAGCCGATCTCTGAGCGTGTCTCTCGGGTATACCTTCAAGCAC 2063
Db 2188 AAAATGCTTATGAAGACAGCTCACCTTATTCCTATTCCTCAGGAGAACTGTCTTCATG 2247
Qy 2064 AAGATGGTGTACGAGGACACCTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2123
Db 2248 TCGATGAAACCCAGCTCTATGATGATTCGGGTGCCACAACCTCAGACTTTCCGAAACAGA 2307
Qy 2124 TCTATGAGAACCCCGCTGTGATTCGGGTGCCACAACAGCACTTCCGCAACCCG 2183
Db 2308 GGCATGACCGCTTACTGAGGTTTCTAGTTGTGACAGAGACACTGGTGTATTTACGAG 2367
Qy 2184 GGCATGACTGCCCTGCTGAAAGTCTCCAGCTGCCAGAAACACACCGGCGACTACTACGAG 2243
Db 2368 GACAGTTATGAAGATATTTTACGATACTTGTGAGTAAACAAATGCCATTTGAACCAAGA 2427
Qy 2244 GACAGCTACGAGCATCTCCGCTACCTGCTGCTCCAGAGAAACAGCCCATCGAGCCCGC 2303
Db 2428 AGCTTCTCCAGAAATTCAGACACCCCTAGCACTAGGCAAAAGCAATTTAATGCCACC 2484
Qy 2304 TCTTCTCCAAAACCTCCCGCACCCGACGCGCTCAGAGCAGTTCCAGCGCCACC 2360

RESULT 15

LOCUS I63424 9009 bp DNA PAT 26-SEP-1997
DEFINITION Sequence 1 from patent US 5663060.
ACCESSION I63424
NID g2480997
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 9009)
AUTHORS Lollar, J.S. and Runge, M.S.

Result	No.	Score	Query Match	Length	DB	ID	Description	Pred. No.
	1	105	2.4	265	24	AA184901	mt93s12.r1 Soares mous	4.00e-116
c	2	71	1.6	252	13	AA754459	97SN1787 Rice Immature	2.65e-63
c	3	61	1.4	247	13	AA754458	97SN1784 Rice Immature	1.50e-48
	4	56	1.3	252	13	AA754459	97SN1787 Rice Immature	2.22e-41

Cp 2404 GTGCTGGCGGT-CATCTCCGCCTGGTGCGCCTTCAG 2370

RESULT 3
LOCUS AA754458 247 bp mRNA EST 20-JAN-1998
DEFINITION 97SN1784 Rice Immature Seed Lambda ZAPII cDNA Library Oryza sativa
cDNA clone 97SN1784, mRNA sequence.
ACCESSION AA754458
NID g2801164
KEYWORDS EST.
SOURCE rice.
ORGANISM Oryza sativa
Eukaryotae; Viridiplantae; Charophyta/Embryophyta group;
Embryophyta; Tracheophyta; seed plants; Magnoliophyta; Liliopsida;
Poales; Poaceae; Oryza.
1 (bases 1 to 247)
AUTHORS Nahm,B.H., Kim,J.K., Cheong,J.J., Kim,S.I., Hahn,T.R., Moon,E.P.,
Kim,W.T., Kim,W.Y., Yang,M.S., Park,R.D., Sohn,U.I., Kang,K.Y.,
Lee,M.C. and Eun,M.Y.
TITLE Large-scale Sequencing Analysis of ESTs from Rice Immature Seed
JOURNAL Unpublished (1998)
COMMENT Contact: Eun M.Y.
Department of Cytogenetics
National Inst. of Agri. Sci. and Tech, RDA
Suwon, Kyunggido, Korea
Tel: 82 331 290 0301
Fax: 82 331 290 0307
Email: myeun@snu20.asti.re.kr
Submitted by Baek He Nahm, Dept of Biological Science, Myongji
University, Yongin, Korea. 449-728 bhna@m bioserver.myongji.ac.kr
Seq primer: M13 Reverse Primer.

FEATURES
source Location/Qualifiers
1..247
 /organism="Oryza sativa"
 /cultivar="Milyang23"
 /note="Vector: pBluescript SK(+); Site 1: EcoRI; Site 2:
XhoI; Directional cDNA library inserted into lambda ZAPII
vector at 5' end with EcoRI and 3' end with Xho I site."
 /db_xref="taxon:4530"
 /clone="97SN1784"
 /clone_lib="Rice Immature Seed Lambda ZAPII cDNA Library"
 /tissue_type="Immature Seed"
 /dev_stage="5 days after pollination"
 /lab_host="E. coli SOLR"

BASE COUNT 7 a 16 c 21 g 34 t 169 others
ORIGIN

Query Match 1.4%; Score 61; DB 13; Length 247;
Best Local Similarity 15.5%; Pred.No. 1.50e-48;
Matches 39; Conservative 125; Mismatches 81; Indels 7; Gaps 7;

Db 1 HWDCITMTVWRGCCOBAMNKKHTHMTBBCCVVRRVGTTNNKGNKRITTTWNCSDNA 60
::: | : : : : : ::::: | ::::: | : : : : : | : : : : :
Cp 2767 AAGCTCACCATGATGTTGCTTCACTCCGCGCGAATAGTAGCTCCGCTGGTGAAGCTAGGCGGCGAGACCACG 2708
::: | : : : : : ::::: | ::::: | : : : : : | : : : : :

Db 61 HCRTYBVHYARSKYC-GTYBY-SNNVDVTNGG-TGVGKTIVNV-HSGMNRCNSSVY 116
: : : | : : : : : | : : : : : ||| | : : : : : | : : : : :
Cp 2707 T-GCTGTTGACTGCTCGCGCGTACAGGGGCTGGGTGAAGCTCCGCTGGTGAAGTCACTCTG 2649
: : : | : : : : : | : : : : : ||| | : : : : : | : : : : :

Db 117 VBTAYCDYBHYBD-RANEHVDTRCTNRDGYCNVTASDNGTSATKRVCTDYKTDSDCGG 175
: : : | : : : : : | : : : : : ||| | : : : : : | : : : : :
Cp 2648 GAACACACTTTTCACTGGGCGACGCTCCGCTCTCTCTGGCGGGTTTCGACGACGTG 2589
: : : | : : : : : | : : : : : ||| | : : : : : | : : : : :

[illegible]

QY	1450	AGTGGGGACACCCCTGCTGATCATCTTCAAGAACAGCGCCAGAGCGCCCTACACATCT	1509
Db	148	KCTASGHHISTNYDVKSSTNTWCVTBSYDKSMHCYH-CSBBVKYHTKVSTRATRSYTCV	206
QY	1510	ACCCACACGGCATCACGGAGCTGCGCCCTCTACAGCCCGCGCTGCCAAGGCGTGA	1569
Db	207	RKYCVNH-HTKKYVKKYHVVBBCBHDTSKCKTWHHTKHWVTST	250
QY	1570	AGCACTTGAAGCATCTCCCATCTCTGCCGGCGGAGATCTTCAAGT	1614

RESULT	5	W45868	365 bp	mRNA	EST	23-MAY-1996
LOCUS						
DEFINITION		mc8h11.1.r1 Soares mouse embryo NME13.5 14.5 Mus musculus cDNA clone 354981 5' similar to gb:38337 Mouse milk fat globule membrane protein E8 mRNA, complete (MOUSE)!. mRNA sequence.				

ACCESSION W45868
NID g1330640
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Eutheria; Rodentia; Sciurognathi; Myomorpha; Muridae
Murinae; Mus.

REFERENCE	TITLE
1 (Bases 1 to 365)	The WashU-HHMI Mouse EST Project
MARRA, M., HILLIER, L., ALLEN, M., BOWLES, M., DIETRICH, N., DUBUQUE, T., GENSEL, S., KUCABA, T., LACY, M., LE, M., MARTIN, J., MORRIS, M., SCHLELLBERG, K., STEPTOE, M., TAN, F., UNDERWOOD, K., MOORE, B., THEISING, B., WYLIE, T., LEMONN, G., SOARES, B., WILSON, R. and WATERSTON, R.	Unpublished (1996)

Contact: Marra M/Mouse EST Project
WashU-BHMI Mouse EST Project
Washington University School of MedicineP
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LML; contact the
IMAGE Consortium (info@image.lml.gov) for further information.
MGI:226781

```

Seq primer: ETPrimer
High quality sequence stop: 284.
Location/Qualifiers
1. .365
   /organism="Mus musculus"
   /strain="C57BL/6J"
   /note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; Site 1: Not 1; Site 2: Eco RI; 1st strand
was primed with a Not I - oligo(dT) primer 15'
TGTTACCATCTGAGTCGGAGCGCGCGGAAATTTTTTTTTTTTTTTT
T 3' ], on equal amounts of mRNA from 2 13.5dpc and 2
14.5dpc embryos [total RNA provided by Minoru Ko, Waseda
State Univ., from 2 ]; double-stranded cDNA was ligated
Eco RI adaptors (Pharmacia), digested with Not I and
cloned into the Not I and Eco RI sites of the modified
pT73 vector. Library went through one round of
normalization, and was constructed by Bento Soares and
M.Fatima Bonaído."
/db_xref="taxon:10090"
/clone="354981"

```

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143	144	145	146	147	148	149	150	151	152	153	154	155	156	157	158	159	160	161	162	163	164	165	166	167	168	169	170	171	172	173	174	175	176	177	178	179	180	181	182	183	184	185	186	187	188	189	190	191	192	193	194	195	196	197	198	199	200	201	202	203	204	205	206	207	208	209	210	211	212	213	214	215	216	217	218	219	220	221	222	223	224	225	226	227	228	229	230	231	232	233	234	235	236	237	238	239	240	241	242	243	244	245	246	247	248	249	250	251	252	253	254	255	256	257	258	259	260	261	262	263	264	265	266	267	268	269	270	271	272	273	274	275	276	277	278	279	280	281	282	283	284	285	286	287	288	289	290	291	292	293	294	295	296	297	298	299	300	301	302	303	304	305	306	307	308	309	310	311	312	313	314	315	316	317	318	319	320	321	322	323	324	325	326	327	328	329	330	331	332	333	334	335	336	337	338	339	340	341	342	343	344	345	346	347	348	349	350	351	352	353	354	355	356	357	358	359	360	361	362	363	364	365	366	367	368	369	370	371	372	373	374	375	376	377	378	379	380	381	382	383	384	385	386	387	388	389	390	391	392	393	394	395	396	397	398	399	400	401	402	403	404	405	406	407	408	409	410	411	412	413	414	415	416	417	418	419	420	421	422	423	424	425	426	427	428	429	430	431	432	433	434	435	436	437	438	439	440	441	442	443	444	445	446	447	448	449	450	451	452	453	454	455	456	457	458	459	460	461	462	463	464	465	466	467	468	469	470	471	472	473	474	475	476	477	478	479	480	481	482	483	484	485	486	487	488	489	490	491	492	493	494	495	496	497	498	499	500	501	502	503	504	505	506	507	508	509	510	511	512	513	514	515	516	517	518	519	520	521	522	523	52
--	---	---	---	---	---	---	---	---	---	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	----

[illegible]

Contact: Wilson RK
WashU-Merck Est Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810

RESULT	9					
LOCUS	AA512671	613 bp	mRNA	EST	08-JUL-1997	
DEFINITION	rv39g02.r1 Soares mouse mammary gland NbMNG Mus musculus cDNA clone					
	863762 5' similar to gb:M38337 Mouse milk fat globule membrane					
	protein E8 mRNA, complete (MOUSE); mRNA sequence.					
ACCESSION	AA512671					
NID	q2250525					
KEYWORDS	EST.					
SOURCE	house mouse.					
ORGANISM	Mus musculus					

REFERENCE	Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS	1 (bases 1 to 613) Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T., Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M., Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B., Theising, B., Wylie, T., Lemmon, C., Soares, B., Wilson, R. and Waterston, R.
TITLE	The WashU-HHMI Mouse EST Project
JOURNAL	Unpublished (1996)
COMMENT	Contact: Marra M/Mouse EST Project WashU-HHMI Mouse EST Project Washington University School of MedicineP 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108 Tel: 314 286 1800

Email: mouseest@watson.wustl.edu
 This clone is available royalty-free through LML; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:507850
 Seq primer: -28ml3 rev2 ET from Amersham
 High quality sequence stop: 273.
 Location/Qualifiers
 1. 613
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /note="Organ: mammary gland; Vector: pT73D-Pac
 (Pharmacia) with a modified polylinker; Site 1: Not I;
 Site 2: Eco RI; 1st strand cDNA was primed with a Not I -
 oligo(dT) primer (5'

```

origin) primer [3
TGTACCAATCTGAAGTGGGACGGCCGGCAATGTTTTTTTTTTTTTTTTTTTTTTTTTT
T 3']; double-stranded cDNA was ligated to Eco RI
adaptors (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of the modified pT73 vector.
RNA provided by Dr. Minoru Ko, Wayne State Univ. Library
constructed and normalized by Bento Soares and M.Fatima
Bonaldo. "
/db_xref="taxon:10090"
/clone="863762"
/clone_lib="Soares mouse mammary gland NDBMG"
/sex="male"
/tissue_type="mammary gland"
/dev_stage="4 weeks"
/lab_host="DH10B"
<1. ->613
mRNA 151 a 163 c 174 g 125 t
BASE COUNT
ORIGIN
Query Match 1.3%; Score 58; DB 7; Length 613;
Best local similarity 63.3%; Pred. No. 3.14e-44;

```


Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu

This clone is available royalty-free through LNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Possible reversed clone: similarity on wrong strand
Insert length: 1452 Std Error: 0.00
Seq primer: mob.REGA+ET
High quality sequence stop: 136.

FEATURES

source

Location/Qualifiers

1. .328

/organism="Homo sapiens"

/note="Organ: heart; Vector: pT73D (Pharmacia) with a
modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st

strand cDNA was primed with a Not I - oligo(dT) primer [5'

TGCTTACCAATCTGAAGTGGGCGCGCATCTTTTTTTTTTTTTTTT 3']

double-stranded cDNA was size selected, ligated to Eco RI

adapters (Pharmacia), digested with Not I and cloned into

the Not I and Eco RI sites of a modified pT73 vector

(Pharmacia). Library went through one round of

normalization to a Cot = 5. Library constructed by

M.Fatima Bonaldo. This library was constructed from the

same fetus as the fetal lung library, Soares fetal lung

NBH119W."

/db_xref="taxon:9606"

/clone="356829"

/clone_lib="Soares fetal heart NBH119W"

/sex="unknown"

/dev_stage="19 weeks"

/lab_host="DH10B (ampicillin resistant)"

<1. .>328

mRNA

BASE COUNT

ORIGIN

Query Match 1.2%; Score 53; DB 22; Length 328;

Best Local Similarity 60.6%; Pred. No. 3.71e-37;

Matches 132; Conservative 0; Mismatches 86; Indels 0; Gaps 0;

Db 106 GGAGTTTNGGTACTGCACAAAACCGGTGCATGTCACCTGTTTGACAGCCCTGT 165

|| ||| | || ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| |

Qy 3815 GGTGTTCTTGGCAAGCTGGACAGCGGCATCAAGCAACATCTTCAACCCGCCAT 3874

Db 166 GGAGCTCAGTACGATGTATGCCACAGAGCTGCCACAGCGCTGCACCTCTGCGCTT 225

|| ||| | || ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| |

Qy 3875 CATCGCCGCTACATCCCGCTGCACCCACCCACTACAGCATCCGACGACCCCTGCCAT 3934

Db 226 TCAGCTACTGGGCTGAGCNGAACGATGCGCAATCCCGGGCTCGAAGTAACAG 285

||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy 3935 GGAGCTGATGGCTCGACCTGACAGCTGCAGCATGCCCTTGGGCATGGAGCAAGGC 3994

Db 286 CATCCCTCAGACAGATCAGCGCCNCCANCANTACA 323

||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy 3995 CATAGGAGCCGACATCACCGGCTTCCAGTACTTCA 4032

RESULT 14

LOCUS H06913

DEFINITION y184d03.r1 Homo sapiens cDNA clone 44854 5' similar to SP:A47285

A47285 MILK FAT GLOBULE PROTEIN - HUMAN ;

ACCESSION H06913

NID g870445

KEYWORDS EST.

401 bp mRNA EST 21-JUN-1995

y184d03.r1 Homo sapiens cDNA clone 44854 5' similar to SP:A47285

A47285 MILK FAT GLOBULE PROTEIN - HUMAN ;

H06913

g870445

EST.

SOURCE human clone=44854 library=Soares infant brain 1N1B vector=Lafmid BA
host=DH10B (ampicillin resistant) primer=M13RP1 Rsite1=Not I
Rsite2=Hind III Whole brain from a 73 days post natal female. 1st
strand cDNA was primed with a Not I - oligo(dT) primer [5'
AACTGGAGAAATTCGGCGCGGAGGAATTTTTTTTTTTTTTTT 3']; double-stranded
cDNA was ligated to Hind III adaptors (Pharmacia), digested with
Not I and directionally cloned into the Not I and Hind III sites of
the Lafmid BA vector. Library went through one round of
normalization. Library constructed by Bento Soares and M.Fatima
Bonaldo.

ORGANISM

Homo sapiens

Eukaryotae; Metazoa; Eumetazoa; Bilateria; Coelomata;
Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes;
Sarcopterygii; Chonata; Tetrapoda; Amniota; Mammalia; Theria;
Eutheria; Archonta; Primates; Catarrhini; Homidae; Homo.

I (bases 1 to 401)

REFERENCE

AUTHORS

Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
Trevaskis,E., Waterston,R., Williamson,A., Wohlmann,P. and
Wilson,R.

The WashU-Merck EST Project

JOURNAL

Unpublished (1995)

COMMENT

GDB: G00-417-395

Contact: Wilson RK

WashU-Merck EST Project

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

High quality sequence stops: 276

Source: IMAGE Consortium, LNL

This clone is available royalty-free through LNL ; contact the

IMAGE Consortium (info@image.llnl.gov) for further information.

FEATURES

source

1. .401

/organism="Homo sapiens"

/clone="44854"

BASE COUNT 83 a 113 c 117 g 85 t 3 others

ORIGIN

Query Match

Best Local Similarity 69.6%; Pred. No. 1.47e-38;

Matches 96; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

Db 37 TGTACCCCGAGGTGCCACGGCGCTGCACTCTCGGCTTGAGCTACTGGGCTGTGACG 96

|| ||||| || || || ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy 3895 TGCACCCACCACTACAGCATCCGACACCTCCGATGGAGCTGATGGCTGGCACC 3954

Db 97 TCAACGGATGCCCAATCCCTGGGCTGAAAGATAACAGCATCCCTGCAAGCAGCATCA 156

||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy 3955 TGAACAGCTGACAGATCCCTGGGCTGGGATGGAGCAGGCCATCAGCCGCCCATCA 4014

Db 157 CGGCTCCAGCAGCTACA 174

| ||||| ||||| |||||

Qy 4015 CGGCTCCAGCTACTTCA 4032

RESULT 15

LOCUS

DEFINITION

AA265280 413 bp mRNA EST 20-MAR-1997

mx91h10.r1 Soares mouse NHL Mus musculus cDNA clone 693763 5'

similar to gb:M16967 COAGULATION FACTOR V PRECURSOR (HUMAN) ; mRNA

ACCESSION	AA265280	sequence.
NID	g1901385	
KEYWORDS	EST.	
SOURCE	house mouse	
ORGANISM	Mus musculus	

REFERENCE
AUTHORS

1 (bases 1 to 413)
Marra,M., Hillier,D., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
Gisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
Waterston,K.
The WashU-RHMI Mouse EST Project
Unpublished (1996)

TITLE	JOURNAL	COMMENT
1. The Role of the Teacher in the Classroom	Journal of Educational Research	1980, Vol. 83, No. 1, pp. 1-10
2. The Impact of Technology on Education	Journal of Educational Research	1980, Vol. 83, No. 2, pp. 11-20
3. The Importance of Parental Involvement	Journal of Educational Research	1980, Vol. 83, No. 3, pp. 21-30
4. The Effect of Teacher Expectations on Student Achievement	Journal of Educational Research	1980, Vol. 83, No. 4, pp. 31-40
5. The Role of the School in the Community	Journal of Educational Research	1980, Vol. 83, No. 5, pp. 41-50
6. The Impact of Teacher Education on Classroom Practice	Journal of Educational Research	1980, Vol. 83, No. 6, pp. 51-60
7. The Importance of Teacher Collaboration	Journal of Educational Research	1980, Vol. 83, No. 7, pp. 61-70
8. The Effect of Teacher Attitudes on Student Behavior	Journal of Educational Research	1980, Vol. 83, No. 8, pp. 71-80
9. The Role of the School in the Development of the Child	Journal of Educational Research	1980, Vol. 83, No. 9, pp. 81-90
10. The Impact of Teacher Training on Student Learning	Journal of Educational Research	1980, Vol. 83, No. 10, pp. 91-100

Contact: Marra M/Mouse EST Project
WashU-HMMI Mouse EST Project
Washington University School of MedicineP
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
MGI:427323
Trace considered overall poor quality
Seq primer: -28m13 rev2 ET from Amer sham
High quality sequence stop: 1.

Db 150 GTAAGTCTCTGTCCTCTGAGATGTACGTGAAG 181
 Qy 4186 TCAAGACCTCTGCTACCAAGCATGTACGTGAAG 4217

Search completed: Thu Jul 9 19:28:46 1998
Job time : 8557 secs.

Result No.	Query %			Length	DB	ID	Description	Pred. No.
	Score	Match						
1	2165	48.6	9354	37	T73164		cDNA encoding human B	0.00e+00
2	2111	47.4	4629	13	Q76016		B-domain deleted Fact	0.00e+00
3	1643	36.9	4275	1	N90654		DNA encoding 740 Arg-	0.00e+00
4	1642	36.9	4275	1	N80446		Modified factor VIII:	0.00e+00
5	1640	36.8	4272	1	N80447		Modified factor VIII:	0.00e+00

File	Key	Location/Qualifiers
FT	CDS	2965..7380
FT		/*tag= a
FT	misc_feature	5165..5174
FT		/*tag= b
FT		/note= '5', near con
FT	misc_feature	5695..5703
FT		/*tag= c
FT		/note= '5', near con
FT	misc feature	6320..6328

RESULT 2
ID Q76016 standard; cDNA; 4629 BP.
AC Q76016;
DT 20-JUL-1995 (first entry)
DE B-domain deleted Factor-VIII.
KW Factor-VIII; blood-clotting; hemophilia A; gene therapy;
KW adenovirus; vector; ss.
OS Homo sapiens.
PN M09429471-A.
PD 22-DEC-1994.
PF 13-APR-1994; 004075.
PR 10-JUN-1993; US-074920.
PR 25-MAR-1994; US-218335.
PA (GENE-) GENETIC THERAPY INC.
PI Connelly S, Kaleko M, Smith T;
DR WPI; 95-036495/05.
DR P-PSDB; R67709.
PT New adenoviral vectors for treatment of haemophilia - contg. a
PT DNA sequence encoding a clotting factor, partic. Factor VIII or
PT Factor IX
PS Disclosure; Fig. 17A-17C; 116pp; English.
CC Human Factor-VIII cDNA, from which the B domain had been deleted, was
CC used to construct recombinant adenovirus vectors that produced
CC therapeutic levels of the clotting factor when administered to an
CC animal host, potentially providing hemophilia A gene therapy.
SQ Sequence 4629 BP; 1318 A; 1051 C; 1018 G; 1242 T;
Query Match 47.4%; Score 2111; DB 13; Length 4629;
Best Local Similarity 74.2%; Pred. No. 0.00e+00;
Matches 3235; Conservative 0; Mismatches 1124; Indels 0; Gaps 0;
Db 58 gccaccagaagatactacctgggtgcagtggaactgtcatgggactatatcaaaagtgat 117
Qy 84 GCCACCGCGCTACTACTTGGCGCGGTGGAGTGTCTGGGACTATACATGCAGAGCGAC 143
Db 118 ctgcgtgagctgctgtggaacgaatttctctctagatgagcaaaatcttttccattg 177
Qy 144 CTGGCGAGCTCCGCTGGACGCGCTTCCGCCCGCGGTGCCAGAGAGCTTCCCGCTC 203
Db 178 aacacctcagtcgtacaaaagactctgtttgtagaattoacgggttcaccttttcaac 237
Qy 204 AACACCGCTGTGTACAGAAACCCCTGTCTGGAGTTCCACCGACCACTGTTCAAC 263
Db 238 atcgttaagcgaagccaccctggatgggtctgctaggtcctaccatccaggctgagtt 297
Qy 264 ATTGCCAAAGCGCGCCCGCTGGATGGGCTGTGGGCCCAACCATCCAGGCCGAGTG 323
Db 298 tatgatacagtgctcatcacacttaagaacatggcttcccatctctgcagctctcatgct 357
Qy 324 TAGCAGCCGTGTGATCACCCTGAAACATATGCCAGCCCGCTGACGCTGCAGCC 383
Db 358 gtcggtgtatcctactgaaagcttctgagggagctgaatatgatgatcagaccagtc 417
Qy 384 GTGGCGTGTGACTTACAGAGCCGACGAGGCGCGCGAGTACGACGACGACGCTCCAG 443
Db 418 agggagaagaagatgataaagttcttctcctggtggaagccatacatatgtctggcagtc 477
Qy 444 CGCGAGAGGAGGACGACAGAGTGTCCCGGGGGGAGCCACCACTAGCTGTGGCAGTG 503
Db 478 ctgaaagagaatggtccaatggctctgaccactgtgccttaccctactactatcttct 537
Qy 504 CTTAAGGAGAACGGCCCTATGGCGAGCGACCCCGCTGTGCTGCTACACTTACCTGAGC 563

Db 538 catgtggacctggtaaaagacttgaattcaggctcattggagccctactactatgtaga 597
Qy 564 CACGTGCACTGTGTAAGGATCTGAACAGCGGGCTGATCGGGCCCTCTGCTGTGTGCGC 623
Db 598 gaaggagctgcccgaaggaaagacacagaccttgacaaaatttatactacttttgc 657
Qy 624 GAGGGCAGCTGGCCAAAGGAAACCAGACCCCTGCACAAAGTTCACTGCTGTGTCGCC 683
Db 658 gtatttgatgaagggaagttggcactcagaacaaagaactccttgatgcaggatagg 717
Qy 684 GTGTTCACAGGGGAAGAGCTGGCAGCGGAGACTTAGAACAAGCTGATCGAGGACGCG 743
Db 718 gatgtcactgtctcgggcctggcctaaaatgcacacagtcgaatgggttatgtaaacagg 777
Qy 744 GACGGCCGAGCGCCCGCGCTGGCCCAAGATGCACACCGTTAACGGGTAGCTGAACCGC 803
Db 778 tctctgccaggtctgattggatgccacaggaatcagtcattggcatgtgattggaatg 837
Qy 804 AGCCTGCGCGGCTGATCGGCTGCCACGCAAGAGCGGTGACTGGCAGCTCATCGCATG 863
Db 838 ggcacaactcctgaagtgaactcaatatctcctgaaggtcacacatttcttggaggaac 897
Qy 864 GGCACCAACCCCTGAGGTGCACAGCATCTTCTCGAGGGCCACACCTTCTGTCGCGAAC 923
Db 898 catcgccaggcgtccttggaaatcgcgaataacttctcctactgctcaaacactcttg 957
Qy 924 CACGCCAGCGCCAGCTGGAGATCAGGCCCATCACTTCTGACTGCCCGACGACCTGCTG 983
Db 958 atggaccttggacagtttctactgttttgtcatatctcttccacacacatgatggcgt 1017
Qy 984 ATGCACTAGCGCAGCTTCTGCTGTTCGCCACATCAGCAGCCACGACGACGCGCATG 1043
Db 1018 gaagcttatgtcaagtagacagctgtccagaggaaccccaactacgaatgaaataaat 1077
Qy 1044 GAGGCTTACCTGAAAGGTGCAGAGCTGCCCGCAGGAGCGCCAGCTGCGCATGAAGAACAAC 1103
Db 1078 gaagaacgggaagactatgatgatcttactgatctcgaaatggatgtggtcagttt 1137
Qy 1104 GAGGAGGCGGAGCACTACGACGACGACCTGACCGACAGCAGATGATGCTGACGCTTC 1163
Db 1138 gatgatgacaactctccttcttccaaattcgtcagttgccaagaagatcctcaaa 1197
Qy 1164 GAGGACGACACAGCGCCACGCTTCATCCAGATCCGACGCTGGCCAAAGACGACCTTAAG 1223
Db 1198 acttggtgacattacattgctgtaagaggagactgggactatgctcccttagtcctc 1257
Qy 1224 ACTTGGGTGCATACATCGCCCGCAGGAGGAGGAGCTGGGACTACGCCCGCTAGTACTG 1283
Db 1258 gcccccgatgcagaagttataaaagtcaaatatttgaacaatggcctcagcgatgggt 1317
Qy 1284 GCCCCGACGACGCGCATACAGAGAGCGAGTACTTGAACACAGGCCCCCGCCGCTCGGC 1343
Db 1318 aggaagtacaaaagtcctgatttatggcatcacagatgaaaccttttaagactcgtgaa 1377
Qy 1344 CGCAAGTACAAAGAGGTGCGCTTCA TGGCCTACACCGACGAGACTTTCAAGACCGCGAG 1403
Db 1378 gctattcagcatgaatcaggaatcttgggacctttactttatggggaagttgagacaca 1437
Qy 1404 GCCATCCAGCAGCACTCGGCACTCTCGGCCCTCTGCTGTACGGCGAGGTGGCGACAC 1463
Db 1438 ctgttgattattttaagaatcaagcagcagcaccataataacatctaccctcaggaatc 1497
Qy 1464 CTGTGATCATCTTCAAGAACAGCGGACGCGCGCTTACACACATCTACCCCGCGCATC 1523

Db	3418	cttattgagcagcatctacatgctgggatgagcacacttttctctgggtacagacaataag	3417
Qy	3444	CTGATCGCGGAGCACTGCACGGCGCATGAGCACCTGTCTCTGTTGTTACGACCAAG	3503
Db	3478	tgtcagactccccctgggaagtgctctggacacacttagagatttttcagattacacgtcca	3537
Qy	3504	TGCCAGACCCCCCTGGGATGGCCACGGGCCACATCCGGCATCTCCAGATCACCGCCAG	3563
Db	3538	ggacaatatggacgtggggccccaaaagctggccagacttcattatccggatcaaatcaat	3597
Qy	3564	GGCAGTATGGGCCATGGGCTCCAAAGCTGGCGCCCTGCATACAGGGCGACATCAAC	3623
Db	3598	gcctcaggacaccaagagacccttttcttgatcaagtggtatgctgtggccaccaatgatt	3657
Qy	3624	GCCTGGTCGACCAAGGACGCCCTCTCTGGATCAAGGTGGACCTGCTGGCCCCCATGATC	3683
Db	3658	attcaaggcaatcaagacccagggtgcctgcagaagttctccagcctctacatctctcag	3717
Qy	3684	ATCACCGCATCAAGACCCAGGGCGCCCGAGGAAGTTAGCAGAGCTGTACATCAGCCAG	3743
Db	3718	tttatcatcatgtatagctcttgatgggaagaagtggcagactctcagagaaattccact	3777
Qy	3744	TTCATCATCATGTACTCTTAGAGCGCAAGAAGTGCAGACCTACCGGGCAACAGCACCC	3803
Db	3778	ggaaccttaatggtctcttttgccaatggtgattcatctgggataaaacacaattttt	3837
Qy	3804	GGCACCTGATGCTTCTCTGGCAGCTGGACAGCAGCGGCATCAAGCACACATCTTC	3863
Db	3838	aacctcccaattattgctcgatacatccgtttgacccaactcattatagcattccgacg	3897
Qy	3864	AACCCCCCATCATCGCCCGGTATCTCCCTGCACCCCACTACAGCATCGATCGCAGC	3923
Db	3898	actcttcgcagtgagttgatgggtgatttaaatagttgcagcatgccattgggaatg	3957
Qy	3924	ACCTTCGCGATGAGGCTCATGGCTCGCACTGAAAGTGCAGCATGCCCTGGCGATG	3983
Db	3958	gagagtaaaaccaatatcagatgcacagattactcttcactctacttaccaatatgttt	4017
Qy	3984	GAGAGGAAGGCCATCAGCGACGCCCGCAGATCACCGCCTCCAGCTACTTCACCAATGTTC	4043
Db	4018	gccacctggtctctttcaaaagctcgaattcaactccaaggagagagtaatgctcgaga	4077
Qy	4044	GCCACTTGGAGCCCAAGAGCCCGCTGCACCTGCAGCGCGCAGCAACGGCTGGCGC	4103
Db	4078	cctcaggtgaataatccaaaagatgggtgcgaatggacttccagaagacaatgaaagtc	4137
Qy	4104	CCCCAGTGAACAACCCCAGAGTGGCTGCAGTGGACTTCAGAAACCATGAAGGTG	4163
Db	4138	acaggagtaactactcaggggataaactctcgttaccagcatgtatgtgaaggattc	4197
Qy	4164	ACTGGCTGACCAACCGAGGGCTCAGACCTGCTGTGACAGCATGTACGTGAGAGTTC	4223
Db	4198	ctcatctccagcagtcaagatggccatcagtggaactcttttttcagaatggcaaatga	4257
Qy	4224	CTGATACAGACAGCCAGACGGCCACCAAGTGGACCTGTCTTCCAAACGGCAAGTGTG	4283
Db	4258	aaggtttttcagggaatatcaagactccttcacactcgtggtgaactctctagacccacg	4317
Qy	4284	AAAGTGTTCAGGGCAACCAAGGACGCTTCAACCGGTGTGAACAGCGCTGGACCGCCGC	4343
Db	4318	ttactgactcgtaaccttcgaattccaccccaaggttggtggcaccaagattgcctcgagg	4377
Qy	4344	CTGTCTGACCGGTACTCTGGCATCCACCCCAAGAGTGGGTGCAACGATCGCCTGGCG	4403

Db	4378	atggaggtctcggtcgcaggccacagacctctactga	4416
Qy	4404	ATGAGGTCTGGCTGCGAGGCCAGGACCTGTACTGA	4442
RESULT	3		
ID	N90654	standard; DNA; 4275 BP.	
AC	N90654;		
DT	26-JUN-1990	(first entry)	
DE	DNA encoding 740 Arg-1649 Glu human Factor VIII:C		
KE	Human Factor VIII:C; Ad.RE.neo; 740 Arg-1649 Glu Factor VIII:C;		
KW	haemophilia A.		
KX			
OS	Homo sapiens.		
FH	Key	Location/Qualifiers	
FT	cds	1..4275	
FT		/*tag= a	
PD	EP-306968-A.		
PN	15-MAR-1989.		
PR	09-SEP-1988; 114769.		
PR	08-APR-1988; JP-085454.		
PA	(KAGA) Chemo-Sero-Therap (Tei.j).		
PI	Sugiyama T, Masuda K, Tajima Y, Yonemura H;		
PI	WPI; 89-078467/11.		
DR	P-PSDB; P91165.		
PT	Prodn. of recombinant human Factor-VIII-C -		
PT	using animal cells transformed with a vector contg. the gene for		
PT	Factor VIII:C and a promoter		
PS	Figure 1(1) - 1(13); 32pp; English.		
CC	When translated, Arg-740 of the carboxyl terminus of the H chain is		
CC	directly bonded by a peptide bond to Glu-1649 of the amino terminus of		
CC	L chain. It is used to transform animal cells so that they produce		
CC	human Factor VIII:C. A pretd. expression vector is plasmid Ad.RE.neo.		
CC	The expression vector has at least one promoter upstream of n90654.		
CC	The transformants can constantly and continuously produce human Factor		
CC	VIII:C in high yield on a commercial scale. The human Factor VIII:C so		
CC	produced is considered to corresp. to the smallest species of active and		
CC	intact Factor VIII:C molecules in the human blood plasma. It is useful		
CC	for treating haemophilia A patients.		
SQ	Sequence	4275 BP; 1245 A; 941 C; 945 G; 1144 T;	
Query Match	36.9%;	Score 1643; DB 1; Length 4275;	
Best Local Similarity	73.7%;	Pred. No. 0.00e+00;	
Matches	3211; Conservative	0; Mismatches 1064; Indels 84; Gaps 23;	
Db	1	gccaccagaagatactacctgggtgcagtggaattctctctagatgccaaatcttttcac	60
Qy	84	GCCACCGCGCGTACTACCTGGCGCGGTGGAGCTCTCTGGAGCTACATGCAGACGCAC	143
Db	61	ctcggtgagctgcctgtggacgcgaagattctctctagatgccaaatcttttcac	120
Qy	144	CTGGGGGACGCTCCCGTGGACGCCCGCTTCCCCCGCGCTGCCAAGACTTCCCTTC	203
Db	121	aacacctcagtcgctgtacaaaaagactctgtttgataattcacgcatcaccttttcac	180
Qy	204	AACACAGCGGTGTGTACAGMAAACCTGTTCTGTGTGAGTTACCGACCACTGTTCAC	263
Db	181	atcgctaagccaaggccacscctgatatggtctgctaggtcctaccatcaggctgagtt	240
Qy	264	ATTGCCAAGCGCGCGCCCCCTGTGATGGCGCTGTGTGGCGCCCAACATCCAGGCCAGGTG	323
Db	241	tatgatacagtggttcattacctaaagaacatggcttccactctgctcagttcttctgct	300
Qy	324	TACGACACCGGTGTGTATCACTCTGAAGCATGGCAGCCACCGCTGCACGCC	383

Db 2214 -----acc--a--a---g--agaa--a--t--a-----actc-- 2230
Qy TCCTTCTCCAAAACCTCCGGCACCACGACGGGTGACGAAGACAGTTCAACGCCACCC 2363

Db 2231 ----gtact--a----ctctt---C--ag--tca---g----a---t--caa----ga----gga- 2258
Qy 2364 CCGCTGCTGAAGGCCACACGAGGGGAGATCACCGGCACACCTCTGCAAAAGCAGCAGGAG 2423

Db 2259 -a--attgatgatgataccatcatcagttgaaatgaagaagaagattttgacatttat 2316
Qy 2424 GAGATCGACTACGACGACACCATCAGCGCTGGAGATGAAGAGGAGACTTCGACATCTAC 2483

Db 2317 gatgagatgaaatcagagcccgagctttcaaaagaaacacgacactattttatt 2376
Qy 2484 GAGAGGACGAGAAACGAGAGCCCGCGCTCTTCCAAAGAAACCCGCCACTACTTTCATC 2543

Db 2377 gctgcaagtggagagctctgggattatgggatgagtagctcccaatgttctaagaac 2436
Qy 2544 GCCGCGGTGGAGCGCTGTGGGACTAGCGCATGAGCAGAGCCGCCACGCTCTGGCGAAC 2603

Db 2437 agggctcagagtgggcagtgccctcagttcaagaagattgtttccagggaatttactgat 2496
Qy 2604 CGCGCCACAGCGGCAGCGCTGCCCACTTCAAGAAGGTGCTTCCAGGAGTTTCACCGAC 2663

Db 2497 ggtctcttactcagcccttataccgtggagaactaaatgaacatttgggactcctgggg 2556
Qy 2664 GGAGGTTTACCCAGGCCCTGTACCGGGGAGCTGACGAGCATCTGGGCCCTGTCTGGC 2723

Db 2557 ccatataagagcagaagttgaagataatcatggttaactttcagaatacaggcctct 2616
Qy 2724 CCCTACTCCGCGCGGAGGTGGAGGACAACATCATGTGTGACCTTCCGGAACCAAGCTCC 2783

Db 2617 cgtccctattctcttattctagccttattcttatgaggaagatcagaggaagagca 2676
Qy 2784 CGGCGCTACTCTTACTCTCTCCCTCATCAGCTACGAGGAGGACAGCGCCAGSGCGCC 2843

Db 2677 gaacctagaaaaactttgtcaagcctaatagaacaaacttacttttggaaagtga 2736
Qy 2844 GAGCCCGCAAGAACTTCTGAAAGCCCAACGAGACTAAGACCTACTTCTGGAAGGTGCAG 2903

Db 2737 catcatatggcaccactactaaagatgagttgactgcaaaagcctgggttattctctgat 2796
Qy 2904 CACCACATGGCCCCACAGGACGAGTTGCACTGCAAGGCCTGGGGCTACTTTCAGCGAC 2963

Db 2797 gttgaacctggaaaaagatgtgcactcaggcctgattggacccttctggtctgccacact 2856
Qy 2964 GTGACATGGGAAGAGACGTGACGCGGCTTGTTCGGGCCCTCTGTGTGTGGCACACC 3023

Db 2857 aacacactgaacctgctcatgagagacaagtgcacatcacaggaatttctgtttctc 2916
Qy 3024 AACACCTGAACCCGCCACCGGAGCGAGGTGACTGTGCAGGAATTTGGCCCTGTTCTTC 3083

Db 2917 accatcttgatgagaccacaaagtgtgacttctcactgaaatatggaagaaactgcagg 2976
Qy 3084 ACCATCTTCAGCAGACTAAGAGCTGTACTTTCACCGAACAATGGAGCAACTGCCGC 3143

Db 2977 gctccctgcaatataccagatggaagatcccacttttaaagagaattatcgcttccatgca 3036
Qy 3144 GCCCCTTCACATTCAGATGGAGATCCCACTTTCAGGAGGACTTACCCCTTCCACGCC 3203

Db 3037 atcaatggctacataatgatgatacactacctggcttagtaatggctcaggatcaaaagatt 3096
Qy 3204 ATCAAGCGCTACATCATGACACACCTCGCCGCGCTTGGTGATGGCCGAGGACAGCGCATC 3263

Db 3097 cgaatgttatctgctcagcatggggcagcaatgaaacacatccattctattcatttcagtgga 3156
Qy 3264 CGCTGTTACCTGCTCTATGGCGACCAACGACAGATCCACAGCATCCACTTCAGCGGC 3323

Db 3157 catgtgttcactgtacgaaaaaaggaggtataaaatggcactgtacaactctctatca 3216
Qy 3324 CACGTTTTTACCGCTGCCAAGAGGAGGAGGTACAGATGGCCCTGTACAACTGTACCCC 3383

Db 3217 ggtgtttttgagacagtggaaatgttaccatccaagctggaatttggcgggtggaatgc 3276
Qy 3384 GCGGTGTTTCAAGACTGTGGAGATGCTGCCAGAGGGCGGGATCTGGCGCTGGAGTGC 3443

Db 3277 cttaatggcagcatctacatgctgggatgagcacacttttctggtgacagacaataag 3336
Qy 3444 CTGATCGGGCAGACCTGTCAGCGCGGCATGAGCACCCCTGTTCTCTGGTGTACAGCAACAG 3503

Db 3337 tctcagactccctgggaatggcttctggacacattagagattttcagattacagcttca 3396
Qy 3504 TGCCAGACCCCTCTGGCATTGGCCATGGCCAGCGGCCACATCCGCGACTTCCAGATCACCGCCAGC 3563

Db 3397 ggaacaataggaacagtgggcccacaagctggccagacttcatttccgggatcaatcaat 3456
Qy 3564 GGCAGTACGGCCAGCTGGGCTCCCAAGCTGGCCGCCCTGCACTACAGCGCGGACATCAAC 3623

Db 3457 gcttgagaccaccaagagccctttcttggatcaagggtgatctgttggaccaatgatt 3516
Qy 3624 GCGTGTGACCAAGAGGCCCTTCTCTGGATCAAGGTGGACCTGCTGGCCCCCATGATC 3683

Db 3517 attcagcgaatcaagaccaggggtgccctcagaagtttccagctctacatctctcag 3576
Qy 3684 ATCCAGCGCATCAACAGCCAGGGCGCGCCCAAGATTCAGCAGCGCTGTACATCAGCCAG 3743

Db 3577 ttatcatcatgtatgtcttgatgggaagatggcagacttatcgaggaattccact 3636
Qy 3744 TTCTCATCATCTACTCTCTTACAGCGCAAGAGTGGCAGACTTACCGCGGCAACAGCAC 3803

Db 3637 ggaacctaatggctcttcttgggaatgtggatctcatctgggaalaaacacaataatttt 3696
Qy 3804 GGCACCTGTATGGTCTTCTTGGCAACCTGACAGCGCATCAAGCACAACATCTTTC 3863

Db 3697 aacctcccaattattgtctgatacatccgtttgcaaccaactcattatagcattcgagc 3756
Qy 3864 AACCCCCCATCATCGCCCGCTTACATCGGCTGACAGCCGCCACTTACAGCATCCCGCAGC 3923

Db 3757 actcttcgcatggagtgtggtgattttaaagtgtgcagcatgccattggggaatg 3816
Qy 3924 ACCCTGCGCATGAGCTGTATGGGCTGCGACCTGACAGCTGCGAGCATGCCCTTGGGCATG 3983

Db 3817 gaaqatlaaacaatatcagatgcacagattactggttcatctacttaccacaatagtgt 3876
Qy 3984 GAGAGAAAGGCCATCAGCGAGCGCCAGATCACCGCTCCAGCTACTTTCACCAACATGTTTC 4043

Db 3877 gcaacctggtctcttcaaaagctcgacttcaectccaaggggaggagtaatgacctggaga 3936
Qy 4044 GCCACTTGGAGGCCCAAGCAAGCGCGCTGCACTGCGAGGGCGCGCAGCAAGCGCTGGCGC 4103

Db 3937 cctcaggtgaataatccaaaagagtggctgcaagtggaactccagaagacaatgaaagtc 3996
Qy 4104 CCCAGGTGACAAACCCCAAGAGCTGGCTGCAAGTGGACTTCCAGAAACCATGAGGTC 4163

Db 3997 acaggagttaactactcaggaggtaaaaactctctgcttaccagcatgtatgtgaagagttc 4056
Qy 4164 ACTGGCGTGACACCCAGCGCGCTCAAGAGCGCTCTGACGACGATGTACGTGAAGGAGTTC 4223

Db	2980	gctccctgcaatattccagatggaagatccccatttttaagagagaattatgcttccatgca	3039
Qy	3144	gccccctgaacattccagatggaaagatccacacttcaaggacmaactacggctttccagcc	3203
Db	3040	atcaattggtcatcaatgaatggtacactacctgggttagtgaatggctcagatcaaaagatt	3099
Qy	3204	ATCAACGGCTACATCATGACACACCTGCCGGCCTGGTGTATGGCCGAGCACCGGCATC	3263
Db	3100	cgatggtatctgctcagcatgggcagcaaatgaaacatccattctattcatttcagtga	3159
Qy	3264	CGCTGGTACTGTCTCTATGGCGACGACAGAAACATCCACGCATCCACTTCAGCGGC	3323
Db	3160	catgtgttcactgtacgaaaaaagaggagtataaaatggcactgtacaatctctatcca	3219
Qy	3324	CAGCTTTTCAACCTGGCCAGAGAGGAGGTACAAAGATGGCCCTGTACAACTGTACCC	3383
Db	3220	gggtttttgaacactggaaatgtlaccatccaaagctggaatttggcgggtggaatgc	3279
Qy	3384	GGCTGTCTCGACACTGTGGAGATGCTGCCCAAGCAAGCCCGGATCTGGCGCTGGAATGC	3443
Db	3280	cttattggcagacatcatcactgctgggtagagacacattttctgtgtacagacaataag	3339
Qy	3444	CTGATCGGCGAGCACTGCACCGCGCATAGCACCCTGTTCTCTGTGTACAGCAACAAG	3503
Db	3340	tgtcagactccctgggaatggcttctggacacattagagatttccagattacagettcca	3399
Qy	3504	TGCCACAGCCCTTGGGCATGGCCAGCGGCCACATCGCGACTTCAGATCAACCGCAGC	3563
Db	3400	ggacaatatggacagtgggcccaaaagctggccagacttcattatccggtacatcaat	3459
Qy	3564	GGCCACTAGGCGCAGTGGGCTCCAGACTTGGCCGGCTGCATACAGCGCGAGCATCMAC	3623
Db	3460	gcttggagcccaaggagccctttcttgatcaaggtggatctgttggccacaatgatt	3519
Qy	3624	GCCTGGTCGACCAAGAGGCCCTTCTCTGGATCAAGGTGGACTGCTTGGCCCCATGATC	3683
Db	3520	attcacggcatcaagaccaggctgccctcagaaagttctccagctctacactctccag	3579
Qy	3684	ATCCAGCGCATCAAGACCCAGCGCGCCGCGAGAAAGTTACAGCGCTGTACATCAGCCAG	3743
Db	3580	tttatcatcatgatatgctctgtaggaaagatggcagaacttatcgaagaaattcca	3639
Qy	3744	TTCATCATCATGTACTCTTTAGCGGCAAGAAGTGGAGACCTACCGCGGCAACAGACC	3803
Db	3640	ggaaccttaatggtctttcttggcaatgtggatctcatctgggataaaacacaattttt	3699
Qy	3804	GGCACCGTCTGTGTTCTTCGGCAACCTGGACAGCAGCGGCGCATCAAGCAACAATCTTC	3863
Db	3700	aacctccaattatgctcogatacatccgtttgcaaccaactcattatgacttcgcagc	3759
Qy	3864	AACCCCCCATCATCGCGCGCTACATCGGCTGCACCCCAACCCACTACAGCATCCGCAACG	3923
Db	3760	actcttcgcatggagttgatgggctgtgatttaaatagttgcagcatgcattgggaatg	3819
Qy	3924	ACCTCTCGCATGAGGTGATGGGCTGCGACCTGCAACAGTGCACGATGCCCTTGGGCATG	3983
Db	3820	gagagttaagcaatatcagatgcacagattactgcttcatctactttaccaaatgttt	3879
Qy	3984	GAGAGCAAGCGGCATCAGACGCGCCAGATCACCGCCTTCCAGCTACTTTCACCAACAATGTC	4043
Db	3880	gccacctggctctccttcaaaagctcgacttccactccaagggaggagtaatgcctcgaga	3939
Qy	4044	GCCACTTGGAGCCCGACGAAGGCCCGCTTCGACCTTCGACGGCGCGCAGCAACGCTGGCCG	4103

Db	3940	cctcaggtgataataccaagaagctggctgcagatggagcttcacagaagaatgaagtc	3999
Qy	4104	CCCCAGGTGAACACCCCAAGGAGTGGCTGCAGCTGGCACTTCAGAAACCATGAGGTG	4163
Db	4000	acagagtaactactcaggagagtaaaatctctgttaccagatgtatgtgaagagttc	4059
Qy	4164	ACTGGCTGACCCACCCAGGGCTCAAGACCTGTCACGAGATGTACGTGAAGGAGTTC	4223
Db	4060	ctcatctccagcagctcaagatggccatcagtcagtgactctcttttttcagaatggcaaaagta	4119
Qy	4224	CTGATCAGCAGCACCCAGCAGCGGCCACAGTGGACCTGTTCTTCGAAACCGCAAGGTG	4283
Db	4120	aaggttttcagggaataacaagactccttcacacctggtggaactctctagaccaccgc	4179
Qy	4284	RAGGTGTTCCAGGGCAACCCAGCAGAGCTTCACACCGGTGTCGACAGCTGCACCCCCCC	4343
Db	4180	ttactgaactcactctcgaaattccattccccccagagttgggtgcaccagatgccctgagc	4239
Qy	4344	CTGCTGACCCGCTACCTGGCGATCCACCCCGAGAGCTGGGTGTCACAGATGCCCTCGCG	4403
Db	4240	atgaaggtctctgggtcgagagccagagacctctac	4275
Qy	4404	ATGCGAGTGCTGGCTGCGAGGCCAGGACCTGTAC	4439

RESULT 5

ID N80447 standard; DNA; 4272 BP.

AC N80447;

DT 10-OCT-1990 (first entry)

DE Modified factor VIII:C sequence with the R740-D1658 deletion.

KE Modified factor VIII:C; haemophilia; procoagulant;

KW blood coagulation; RD deletion; ss.

KX Homo sapiens.

PN W08800831-A.

PD 11-FEB-1988.

PR 31-JUL-1987; U01814.

PR 01-AUG-1986; U5-893375.

PA (BIOJ) Biogen NV (PASE/).

PI Pasek MP;

PR WPI; 88-049866/A7.

DR P-PSDB; P80268.

DT New DNA sequences encoding modified factor VIII:C - with deletion of DNA

PT encoding maturation polypeptide, useful for high yield transformation.

PS Claim 3; Page 47-48-49-50; 97pp; English.

CC The RD deletion removes the DNA from Ser 741 to Ser 1657. A major part

CC of the sequence encoding the maturation polypeptide of

CC factor VIII:C is deleted, i.e. Gln 744 - Asp 1563.

CC The full length Factor VIII:C cDNA has two changes with respect to the

CC published sequence (EPO application 160457):

CC CTG to CTA at Leu 242 and TTC to CTC change at amino acid residue 1880

CC (Phe to Leu). The product is produced in approx. 20 times higher

CC yields than previous recombinant produced factor VIII:C and are more

CC easily purified. The peptide is used for treating haemophilia A, both

CC acute and prolonged bleeding.

CC See also N80444 and N80446.

SQ Sequence 4272 BP; 1243 A; 941 C; 944 G; 1144 T;

Query Match 36.8%; Score 1640; DB 1; Length 4272;

Best Local Similarity 73.6%; Pred.No. 0.00e+00;

Matches 3208; Conservative 0; Mismatches 1064; Indels 84; Gaps 23;

Db	1	gccaccagaatactacctgggtgcagtgaaactgcatggggactatatgcgaagtcat	60

QY	3924	ACCTGCGCATGGAGCGCTGATGGGCTGGCAACCTGACACAGCTGCAGCATGCCCCCTGGCGCATG	3983
Db	3817	gagagtaaagcaaatatccagatgcacagattactggttcaatcctcactttaccacaatatggtt	3876
QY	3984	GAGACGAGGCCATCAGCGACGGCCGACATACCGGCTCCAGCTACTTCCACCAACATGTTT	4043
Db	3877	gccacctggtctctcttccaaagctcgacttcaoctccaagggagagtgtaatgcctggaga	3936
QY	4044	GCCACTTGGAGCCCCAGCAAGGCCCGCTTGCACTGCAAGGGCCGACGACGCCTGGCCG	4103
Db	3937	cctcaggtggaataatccaaagagtggtgcgaagtggacttccagagaacataaataaagtc	3996
QY	4104	CCCCAGGTGAACAAACCCACAGGATGGCTGCAGGTGGACTTCCAGAAAAACATGAGGTG	4163
Db	3997	acaggagtaactactcaggagagtaaaaatctctgcttaccagactgtatgtgaagagattc	4056
QY	4164	ACTGGCGCTGACCAACCCAGGGGCTCAAGAGCCGTGTGACGACATGTACTGAGGAGTTT	4223
Db	4057	ctcatctccagcagtcagaatggccatcagtggaactctctttttcagaatggcaagta	4116
QY	4224	CTGATCAGCAGCGCCAGGACGGGCCACAGTGGACCCCTGTTTCCAAAACGGCAAGGTG	4283
Db	4117	aaggtttttcaggaaatacaagactccttcacacctgggtgaactctctagacaccagc	4176
QY	4284	AAGGTGTTCCAGGGGCAACACGACACTTTCACACCGGTGCTGACAGCTTGACCCCCCC	4343
Db	4177	ttactgactcgcctacctctcgaaattcaccocccagagttgggtgcaccagatgccctgaag	4236
QY	4344	CTGCTGACCCGCTACCTGCGCATCCACCCCAAGAGTGGGGTGACAGATGCGCCCTGGCG	4403
Db	4237	atggaggttctgggctgcgaggcaacgaacctctac	4272
QY	4404	ATGGAGTCTTGGGCTGCGAGGCCGAGGACCTGTAC	4439

RESULT

ID	N80444	standard; DNA; 4545 BP.
AC	N80444;	
DT	10-OCT-1990	(first entry)
DE	Modified factor VIII:C sequence with the Q744-D1563 deletion.	
KW	Modified factor VIII:C; haemophilia; procoagulant;	
KW	blood coagulation; QD deletion; ss.	
OS	Homo sapiens.	
PN	W08800831-A.	
PD	11-FEB-1988.	
PF	31-JUL-1987; U01814.	
PR	01-AUG-1986; US-893375.	
PA	(BIO) Biogen NV (PASE/).	
PI	Pasek MP;	
DR	WPI: 88-049866/07.	
DR	P-P5DB; P80265.	
PT	New DNA sequences encoding modified factor VIII:C - with deletion of DNA	
PT	encoding maturation polypeptide, useful for high yield transformation.	
PS	Claim 3; Page 38-39-40-41; 97pp; English.	
CC	A major part of the sequence encoding the maturation polypeptide of	
CC	factor VIII:C is deleted, i.e. Gln 744 - Asp 1563. The QD deletion	
CC	retains approximately 90 amino acids of the maturation polypeptide	
CC	(four amino acids at the N-terminal end and 86 amino acids at	
CC	the C-terminal end). The full length Factor VIII:C cDNA has two	
CC	changes with respect to the published sequence (EPO application 160457):	
CC	CTG to CTA at Leu 242 and TTC to CTC change at amino acid residue 1880	
CC	(Phe to Leu). The product is produced in approx. 20 times higher	
CC	yields than previous recombinant produced factor VIII:C and are more	
CC	easily purified. The peptide is used for treating haemophilia A. both	

Db 844 catcgccaggtccttggaaatctgcgaataactttcttaactgtctcaaacactcttg 903
Qy 924 CACGCCAGGCCAGCGTGGAGATCGCCCAATCACTTCCTGCTGCCAGACCTGCTG 983
Db 904 atggaccttggacagtttctactgtttgtgcataatctcttccaccacaaatgatggcatg 963
Qy 984 ATGACCTTAGGCCAGTTCCTGCTGTTCTGCACATCAGCAGCCACGACGACGCGCATG 1043
Db 964 gaagcttatgtcaaaatagacagctgtccagaggaaccccaactacgaatgaaaaaaat 1023
Qy 1044 GAGGCTTAGCTGAAGTGTGACAGCTGCCCGAGGAGCCCGACGCTGCGCATGAAGAACAAC 1103
Db 1024 gaagaagcggaagactatgatgatctactgtattctgaatggatctgtgaggttt 1083
Qy 1104 GAGGAGGCCGAGGACTTACGACGACGCTTGACCCAGCAGCATGTGATGCTGACGCTTC 1163
Db 1084 gatgatgacaactctctctcttctatccaaattctgcagttggcaagaagcatcctaaa 1143
Qy 1164 GAGCAGCAACAACGCCOAGCTTCATCCAGATCCGCGAGCGTGGCCAAAGAACCTTAAG 1223
Db 1144 acttgggtacattacatgtctgaagagagagactgggactatgctcccttagctctc 1203
Qy 1224 ACCTGGGTGCACATACATCGCCCGCGAGGAGGACTGGGACTAGCGCCCGCTAGTACTG 1283
Db 1204 gccccgatgacgaagttataaagtcaattatttgacacatggccctcagcggtatgggt 1263
Qy 1284 GCCCGGAGCGCGAGCTTACAGAGCCAGTACCTGAACAGCGGCCCGCCGACGCGCATGGC 1343
Db 1264 aggaagtacaaaaagtcgattttatggcatcacagatgaacctttaagactcgtgaa 1323
Qy 1344 CGCACTACAGAGTGGCTTCATGGCTTACACCGACGACACTTTCAGAGCCGCGAG 1403
Db 1324 gctattcagcatgaatcaggaatcttggacctttactttatggggaagtggagacaca 1383
Qy 1404 GCATCCAGACGACGTCGGGATCTCGGCCCTCTGGCCCTGCTGTACGGGAGGTGGCGACAC 1463
Db 1384 ctgttgattatttaagaatcaagaacagacacataaacatcaccctcacggaatc 1443
Qy 1464 CTGCTGATCATCTTCAAGAACCGAGCGCAGCGCCCTTACAAATCTACCCCCACGCGATC 1523
Db 1444 actgatgtccgtcttctgtattcaagagagattaccaaagggtgtaaaactttgaaggt 1503
Qy 1524 ACCGAGTGGCGCCCTGTACAGCGCGCGCTGGCCCAAGGGGTGAAGCACTGTGAAGGAC 1583
Db 1504 ttccaattctgccaggagaatattcaaatataaatggacagtgactgtagaagatggg 1563
Qy 1584 TTCCCATCTCTGCCGCGGAGATCTTCAGTACAGTGGACGCTGACCGCTGGAGGACGGC 1643
Db 1564 ccaactaaatcagatcctcgttgcctgaccctgtattactctagtcttgcgttaatatggag 1623
Qy 1644 CCACCAACAGACGACGCCCGCTGCTGACGCCCTTACTACAGCAGCTTCGTGACATGGAG 1703
Db 1624 agagactagctcagagactcattggccctctcctcatctgctacaaagaatctatagat 1683
Qy 1704 CGGACCTGGGCTTCGAGCTGATCGGCCCTCTGCTGATCTGCTACAGGAGAGCGGTGAC 1763
Db 1684 caaagagaaaccagataatgtcagacagagagaatgtcatcctgtttttctgtatttgat 1743
Qy 1764 CAGCGCGGCAACCAACATCATGAGCGCAACAGCGCAACCTGATCTGCTTCCAGCTGTTCGAC 1823
Db 1744 gagaaccaaagctggtaacctcacagagaatatacaacgctttctcccaatccaagctgga 1803
Qy 1824 GAGACCGGACCTGGTATCTGACCGAGACATCCAGCGCTTCCTGCTCCACACCGCGTGC 1883

Db 1804 gtgcagcttgaggtatccagagttccaagcctccaacatcatgcacagcatcaatggctat 1863
Qy 1884 GTGAGCTGGAAGATCCCGAGTTCAGGCGCAGCAACATCATGCAAGCATCAACGGGTAC 1943
Db 1864 gtttttgatagtttgcagttgtcagtttggatttggcatgaggtggcatactggtacattcta 1923
Qy 1944 GTGTTGCAGAGCTGCAGCTGAGCGCTGTGCTGCATGAGGTGGCTTACTTGGTACATCCTG 2003
Db 1924 agcattggagcacagactgacttctcttctgtctctctctctctgtatatacttcaaacac 1983
Qy 2004 AGCATGGCGGCCACGACGACTTCCTGAGCGTCTTCTCGGGGTATACCTTCAAGCAC 2063
Db 1984 aaaaatggtctatgaagacacactcacctattccattctcaggagaaactgtcttcagt 2043
Qy 2064 AAGATGGTGTACGAGACACCTGACCTGCTCCCTTCTCCGGGAGACTGTGTTCTATG 2123
Db 2044 tcgattgaaaaaccaggtctatgatttctgggtgccacaactcagacttttcggaacaga 2103
Qy 2124 TCTATGAGAACCCCGGCTGTGGAATTCGGGCTGCCAACAGCGACTTCCGCAACCGC 2183
Db 2104 gcatgacgccttactgaaggtttctagtttgcagaagaacactggfagtattactagag 2163
Qy 2184 GGCATGACTGCCCTGCTGAAGTCTCCAGCTGCCAAGAACACACCGGCACTACTACGAG 2243
Db 2164 gacagttatgaagatatctcagcataacttctgagttaaaaaacaatgccattgaaccaaga 2223
Qy 2244 GACAGCTACGAGGACATCCCGCTTACCTGCTGTCCAAGAACAAACGCCAT-----C--GA 2296
Db 2224 agcttctcccagatcctcttctgttgggataaccaactatggttactcagatcaccaaaaga 2283
Qy 2297 -GC-----CC-C--GCT-----C-CT-----T-CTC-----CCAAAC-- 2318
Db 2284 gagtggaaatcccaagagaagtcaaccagaaaaaacagctttttaagaaaaagataccatt 2343
Qy 2319 -----TCCC-----GCCACCC-----CAGC-----ACG-----C----- 2337
Db 2344 ttgtccctgaacgttgtgaaagcaatcatgcaatagcagcaataaatgagggacaaaat 2403
Qy 2338 --GTCA--GAA--GC-----AG--T--T--CAA--C--GC-----CA----- 2358
Db 2404 aagccgaatagaagtcaactcaggtggaagcaagtaggactgaaaggctgtctctcaa 2463
Qy 2359 ---CCC-----C-CC-----C-----GT--G-CTGAA--GC-----CA-- 2380
Db 2464 aaccacacagtttgaacacgcctcaacgvggaaataactcgtactactcttcagtcagat 2523
Qy 2381 ---CCAGC-G-C-GA---G--ATCACCGCGACC--ACCC-TGCAG-----AG-C-GAC 2417
Db 2524 caagagaaattgactatgatgatataccatcagttgaaatgaagaaggaattttgac 2583
Qy 2418 CAGGAGAGATCGACTACGACGACCATCAGCGTGGAGATGAAGAAGAGGAGACTTCGAC 2477
Db 2584 attatgatgaggtgaaaatcagagcccccgagcttcaaaagaagaacacagacat 2643
Qy 2478 ATCTACGACGAGCAGAACACAGAGCCCCCGCTGCTTCCAAAAGAAAACCCGCCACTAC 2537
Db 2644 ttattgctcagtgagaggtctctgggattatggatgagtagtgcctccacatgttcta 2703
Qy 2538 TTATCGCGCGCTGAGCGCTGTGGACTAGCGCATCGACGACGAGCCGCCCGCTCTG 2597
Db 2704 aagaacagggctcaggtggcaggtgctcactcagttcaagaagtgttttccaggaattt 2763
Qy 2598 CGCAACCGCGCCCAAGCGGCGCGCTGCCCGACAGTTCAGAGAGGTGTGTTCCAGGAGTTC 2657

Db 2764 actgaggtccttactcagcccttataccgtggagaaactaaataaacatttggactc 2823
Qy 2658 ACCGACGGCAGCTTCAACCAAGCCCTGTACCGCGGCGAGCTGAACAGCAGCCTGGGCTG 2717
Db 2824 ctggggccatataaagagagaagtgaagataataatcattgtaacttcaagaaatcag 2883
Qy 2718 CTCGGCCCTTACATCCGCGCGAGGTGGAGGACAAATCATGTGTGACTTCGGCAACAA 2777
Db 2884 gctctogtccctattccttctattctagccttatttcttatgaggagaatcagagcaa 2943
Qy 2778 GCCTCGGGCCCTACTCTTCTTACTCTCCCTGATCAGTACGAGGAGNCCAGCGCCAG 2837
Db 2944 ggaagagaactagaaaaactttgtcaagctaatgaaccccaaaacttaacttttgaaa 3003
Qy 2938 GCGCGCGAGCCCGCAGAACTTCTGTGAGCCCAACGAGACTAGACTACTTCTTGGAG 2897
Db 3004 gtgcaacatcatatggcaccactaaagatgagtttgactgcaagacctgggcttatttc 3063
Qy 2998 GTGAGACACCATAGGCCCCCAAGAGCAGGTTGACTGCAAGGCGCTGGGCCCTACTTC 2957
Db 3064 tctgatgtgacctggaaaaagatgtgcaactcaggcctgattgaccccttctggtctgc 3123
Qy 2958 AGCGACGTGGACCTGGAGAGGAGCTGCACAGCGGCTGTATCGCCGCCCTGTGTGTGTC 3017
Db 3124 cacactaacaactgaacctgtctatgggagacaagtgaacagtgaagatttgcctg 3183
Qy 3018 CACACCAACACCTGAACCCCCCGCAGGAGGAGTGACTGTGAGGAATTTGCCCTG 3077
Db 3184 ttttcaccattttgatgagacaaaagctggtaactcaactgaaaatgtgaaagaac 3243
Qy 3078 TTCTTACCATCTTCCAGCAGCTAGAGCTGTGACTTACCGGAGACATGGAGCGCAAC 3137
Db 3244 tgcagggtccttgaatcaatcagatgaagatcccaacttttaagagaattatcgcttc 3303
Qy 3138 TGGCGGCGCCCTGCNAATCCAGATGGAAATCCCAACCTTCCAGGAGAACTACCGCTTC 3197
Db 3304 catgcaatcaatggctacataatggatacactacctggcttagtaatggctcaggatcaa 3363
Qy 3198 CAGCGCATCAACGGCTTACATATGACACCCCTGCGCGCTGTGTATGGCCAGGACGAG 3257
Db 3364 aggattcagtgatctgctcagcagatggcgagcaatgaaacaatccattctatttc 3423
Qy 3258 CGCATCCGCTGGTACCTGCTGTATGGCAGCAACGAGACATCCACGATCCACTTC 3317
Db 3424 agtgagacatgttctactgtacgaaaaaaggaggagtataaaatggcactgtacaatctc 3483
Qy 3318 AGCGGCGAGCTTTTACCGCTCGCAGAGAGGAGGAGTACAAATGATGGCCCTGTATCAACCTG 3377
Db 3484 tatccagggtgttttgagacagtggaatgttaccatccaaagctggaatttggcggtg 3543
Qy 3378 TACCGCGCGGTGTTCAGACTGTGGAGATGTGCGCAGCAAGCGCGGATCTGGCGGTG 3437
Db 3544 gaatgcttattggcagacatctacatgctgggatgagcacacttttctggtgtacagc 3603
Qy 3438 GAGTGCCTGATCGGCGGACCATGCACGCGGCGATGAGCACCTGTTCCTGGGTGTACAGC 3497
Db 3604 aataagtgacagactcccctgggaatggcttctggacacattagagattttcagattaca 3663
Qy 3498 AACAACTGCCAGACCCCTGGCGATGGCAGCGGCCACATCCGCGACTTCCAGATCAC 3557
Db 3664 gcttcaggacaatatgacagatggcccccaagctggccagacttctatttcaggatca 3723
Qy 3558 GCCAGCGGCAGTACGCCAGTGGGTCCCAAGCTGGCCGCTGCATCTACAGCGGCAGC 3617

Db 3724 atcaatgctggagaccagagagcccttttcttgatcaagtggaatctgttggcaca 3783
Qy 3618 ATCAACCGCTTGTGCACCAAGAGAGCCCTTCTCTGGATCAAGGTGAGCTCTGTGGCCCC 3677
Db 3784 atgattattacaggcatcaagaccagggtgcccgtcaagaattctccagcctctacac 3843
Qy 3678 ATGATCATCCAGGTCATCAAGACCCAGCGGCGCCCGCAAGATTTCAGCAGCCTGTATATC 3737
Db 3844 tctcagtttatcatcatgtatagttctgtatgggaagaagtgcagacttctcagagaaat 3903
Qy 3738 AGCCAGTTTCATCATCATGTACTCTTAGACGGCAAGAAAGTGGCAGACCTTCCGCGCAAC 3797
Db 3904 tccactggaacctaatggctcttttggcaatgtggattcctctgggataaaacaaat 3963
Qy 3798 AGCACCAGGACCTTGATGCTGTTCTTCGGCAAGSTGGAGAGCGGCGCATCAAGCAAC 3857
Db 3964 atttttaacctccaattattgctcgatcacatccgtttgaccccaactcattatgactt 4023
Qy 3858 ATCTTCAACCCGCCCATCATTCGCCCGCTTACATCCGCTGCAACCCACCTACAGCATC 3917
Db 4024 cgcagcactctcgcagtgagttgatggctgtgatttaaatagttgcagcatgccattg 4083
Qy 3918 CGCAGCACCTCGCGATGGAGCTGATGGGCTGCGACCTGAACAGCTGCGCATGCCCCCTG 3977
Db 4084 ggaatgagagtaaaagcaatatacagatgcacagattactgcttcactcctactttaccaat 4143
Qy 3978 GGCATGGAGCAAGGCGCATCAGCGACGCCAGATCAGCCGCTCCAGCTTACTTCAACAC 4037
Db 4144 atgtttgccacctggctccttcaaaagctcgacttcactcccaaggaggagtaatgc 4203
Qy 4038 ATGTTGCGCACTTGAGCCCCAGCAAGCGCGGCTGCACCTGCAAGGCGCGCAACGCC 4097
Db 4204 tggagacctcaggtgaataatccaaaagtggtgcaagtggacttccagaagaacatg 4263
Qy 4098 TGGCGCCGCCAGGTGCACACCCCAAGAGTGGCTGCGAGTGGAGCTTCCAGNAACCATG 4157
Db 4264 aaagtcacagagtaactactcagggagtataaatctctgttaccagcatgtatgtgaag 4323
Qy 4158 AAGGTGACTGGCTGACACCCAGGCGCTCAAGCGCTGCTGACAGCATGTACGTGAAG 4217
Db 4324 gaggctcctcatctccagcagtcgaagatggccatcagtggaactctctttttcagaatggc 4383
Qy 4218 GAGTTCTCATCAGCAGCAGGAGCGGCCACCACTGAGCCCTGTCTTCCAAAACGGC 4277
Db 4384 aaagtaagggttttccagggaataatcaagactccttcacacctgtgggaactctctagac 4443
Qy 4278 AAGGTGAAGTGTTCAGGGCAACAGCAGCAGGTTCCACCGCTGCTGACAGCCTGGAC 4337
Db 4444 ccaccgttactgactcgtactctgaattcaccocccagagttgggtgcaccagattgccc 4503
Qy 4338 CCCCCTCTGTCAGCCCGCTACTGCGGATCCACCCCGAGAGCTGGGTGCACCATGCGC 4397
Db 4504 ctgaggaatgaggttctgggctgaggaagcagagacctctac 4545
Qy 4398 CTGCGCATGAGGCTGCTGGGCTGCGAGCGCCAGGACCTGTATC 4439

RESULT 7

ID T03571 standard; cDNA; 6300 BP.

AC T03571;

DT 13-MAY-1996 (first entry)

DE Factor-VIII cDNA.

KW Factor-VIII; blood clotting; haemophilia A; ss.

OS Homo sapiens.

RESULT	ID	T
1	1	1
2	2	2
3	3	3
4	4	4
5	5	5
6	6	6
7	7	7
8	8	8
9	9	9
10	10	10
11	11	11
12	12	12
13	13	13
14	14	14
15	15	15
16	16	16
17	17	17
18	18	18
19	19	19
20	20	20
21	21	21
22	22	22
23	23	23
24	24	24
25	25	25
26	26	26
27	27	27
28	28	28
29	29	29
30	30	30
31	31	31
32	32	32
33	33	33
34	34	34
35	35	35
36	36	36
37	37	37
38	38	38
39	39	39
40	40	40
41	41	41
42	42	42
43	43	43
44	44	44
45	45	45
46	46	46
47	47	47
48	48	48
49	49	49
50	50	50
51	51	51
52	52	52
53	53	53
54	54	54
55	55	55
56	56	56
57	57	57
58	58	58
59	59	59
60	60	60
61	61	61
62	62	62
63	63	63
64	64	64
65	65	65
66	66	66
67	67	67
68	68	68
69	69	69
70	70	70
71	71	71
72	72	72
73	73	73
74	74	74
75	75	75
76	76	76
77	77	77
78	78	78
79	79	79
80	80	80
81	81	81
82	82	82
83	83	83
84	84	84
85	85	85
86	86	86
87	87	87
88	88	88
89	89	89
90	90	90
91	91	91
92	92	92
93	93	93
94	94	94
95	95	95
96	96	96
97	97	97
98	98	98
99	99	99
100	100	100

Query Match 24.4%; Score 1085; DB 33; Length 7053;
Best Local Similarity 73.8%; Pred. No. 0.00e+00;
Matches 1681; Conservative 0; Mismatches 596; Indels 0; Gaps 0;

Db	58	gccaccagaagatactacctgggtggcagtggaactgtctatgggactatattgcaaatgat	117
Qy	84	gccaccggcgctactactggggcggtggagctgtctctgggactatcatgcagagccac	143
Db	118	ctcgtggagctgacctgtggagcgcaagatttctctctatagatgccaaactctttccattc	177
Qy	144	ctggggcagctcccgctggacgcgccgcttccccccggcggtgcccaagactttccgcttc	203
Db	178	acaacctcaagtctgtacaaaaagaactctgtttgtagaattcaagatcacacttttcaac	237
Qy	204	aacacacagcctgggtgtacaaagaaaacctgttctggagtttaccgcacactgtttcaac	263

Db	1144	gatgatgacaactctccttccttataccaatctcgtctcagtttgccaagaagcatcctaa	1203
Qy	1164	GAGCAGCAACACGCCGAGCTTCATCCAGATCCGACGCTGGCCAGAAGACGCCCTAAG	1223
Db	1204	acttgggtacattacattgtctgtgaagagaggaactgggaactatgctcctttagtcctc	1263
Qy	1224	ACCTGGGTGCACATACATCGCCCGCGAGGAGGAGCACTGGGACTAGGCCCGGTAGTACTG	1283
Db	1264	gcoccccgatgacagaagtataaaagtcaatattgaacaatggccctcagcgattggt	1323
Qy	1284	GCCCCGAGCAGCCGAGCTACAGAGCCAGTACCTGAAACAGCGCCGCCGCGCATCGGC	1343
Db	1324	agggaagtacaaaaagtccgattttatggcatcacagatgaaacctttaagactcgtgaa	1383
Qy	1344	CGCAGTTCAAAGAAAGTGGCGTTCATGGCCTACACGACGACGACTTTCGAAGACCCGCGAG	1403
Db	1384	gctattcagcatgaatcaggaactctgggaaccttactttatggggaagtggagacaca	1443
Qy	1404	GCCATCCGACGACGATCCGGCATCTCTGGCGCCCTGCTGACGGCAGGTGGCGCACCC	1463
Db	1444	ctgtgattatattaaagatcaagacaagcagagaccatataaacatctaccctcaggaatc	1503
Qy	1464	CTGCTGATCATCTTCAAGAACGAGCGCAGCAGCGCTACACATCTACCCCGCAGCATC	1523
Db	1504	actgatgtcogtcttctgtattcaaggagattaccaaaagtgtaaacatttgaagat	1563
Qy	1524	ACCGAGGTGGCCGCCCTGTACAGCGCGCCTGCCAGCGGGCTGAAGCATCTGAAGGAC	1583
Db	1564	tttccaattctgccagagagaataattcaatatataatggacagtgactctagaagatggg	1623
Qy	1584	TTCCCATCTCTGCCCGGCGAGATCTTCAGTACAGTGGACCGCTGACCTGGAGGAGCGC	1643
Db	1624	ccaactaaatcagatcctcgttgcctgaccgctattactctagtttctgttaatatggag	1683
Qy	1644	CCCACACAGACGACCCCGCTGCTGACCCGCTACTACGACGCTTGTGTGAACATGGAG	1703
Db	1684	agagatctagcttcaggaactcattggccctctcctcatctgctacaaagaatcttagat	1743
Qy	1704	CGCGACTGGCCTCCGGACTGATCGCGCCCTGCTGATCTGCTACAAGGAGAGCTGGAC	1763
Db	1744	caaaagagaacacagataatgtcagacaagaggaatgtcctcctgtttctctgtattgat	1803
Qy	1764	CAGCGCGGCACACCATATGAGCGACAGGCGCAACGCTGATCTCTTCAAGCTGTTCCGAC	1823
Db	1804	gagaaccgaagctggtacctcacagagaatatacaacgctttctccccatccagatgga	1863
Qy	1824	GAGAACGCGAGCTGGTATCTGACCGAGAACATCCACGGCTCTCTGCCAACCCCGCTGGC	1883
Db	1864	gtcagcttgaggatccagaggttccaagctccaacatcatgcacgatcaaatggtcat	1923
Qy	1884	GTGAGCTGGAGATCCCGAGTTCCAGGCGCAGCAACATCATCGACAGCATCAACGGCTAC	1943
Db	1924	gtttttgatagtttcagttgctcagtttgttgcatgagtggaactactggtacattcta	1983
Qy	1944	GTGTTGCAACCCCTGACGCTGAGCGCTGTCCCTGCAAGAGTGCGGCTACTGTGATCCTCTG	2003
Db	1984	agcattggagcacagactgacttcctttctgtctcttctctcgtgatatacctcaaacac	2043
Qy	2004	AGCATCGCGGCCACGACGACTCTCCTGAGCGGTGTTCTCTCGGGGTATACCTTCAAGCAC	2063
Db	2044	aaaatggtctatgaagacacactcaacctattcccattctcaggaagaactgctctcatg	2103
Qy	2064	AAGATGGGTATACGAGGACACCCCTGACCTTTCCGCTTTCCGCGGAGACTGTGTCTATG	2123

[illegible]

Db	2938	ccaccagctcttgaacggccatcaacgagggaataaactcgtactactcttccagtcagatca	2999
Qy	2360	cccccccgctgctgaaggccacacgaggagatcaccgaccaccccttgcnaaggacca	2419
Db	2998	agagaaattgactatgatgataccatactcgttgaaatgaagaaggaaatttgacat	3057
Qy	2420	ggagagatctgactacacacacacatcaggcgtcgagatcagaagagagacttcgacat	2479
Db	3058	ttatgatgagatgaaatcagagcccccagctttcnaaagaacacagacactattt	3117
Qy	2480	ctacgacgagacagAACACAGAGCCCCGCTCTTCCAAAGAAACCCGCCACTACTT	2539
Db	3118	tattgctgcagtgagagcctctgggattatggatgagtagctccacatgttctaa	3177
Qy	2540	catcgccgcccgtggagagcctctgggacatcagcagcagcccccacgctctccg	2599
Db	3178	aaacaggcctcagagtggcagtgccctcagctcaagaagaattgtttccagaatttca	3237
Qy	2600	caaccggccacagcgcgacgctgcccagcttcAAGAAAGTGGTGTTCAGGAGTTCAC	2659
Db	3238	tgatgctccttactcagcccttataccgtggagaaactaaatgaacatttgggaactc	3297
Qy	2660	cgacggcagcttccacgagccctgtacggcgccagctgaacagcagcactgggctgct	2719
Db	3298	ggggccatatataagagcagaagtgaagataataatcgttaacttccagaatacggc	3357
Qy	2720	cgccgcccatacatccgcccgggtggagacacacATCATGCTGACTTCCGCCAACCA	2779
Db	3358	ctctcgctcctattcctattctagcctatttcttatgaggagatcagagcaagg	3417
Qy	2780	ctccggcccttactccttctactcctcctgATCAGCTACGAGGAGGACACGCCAGG	2839
Db	3418	agcagaacctagaaaactttgcaagcctaatagaacccaaaacttacttttggaaagt	3477
Qy	2840	cgccgagcccccgaagAACCTTCGTGAAGCCCAACAGACTTAAGACTTCTTGAAGGT	2899
Db	3478	gcaacatcataatggcaccactaaagatgattgactgcaagcctgggcttattctc	3537
Qy	2900	gcagcacacatggcccccacgaagcagctgcactgcaaggcctgggcccctacttcag	2959
Db	3538	tgatgtgacctgaaaaagatgtgcactcaggcctgattgacccctcttgctcgcca	3597
Qy	2960	cgacgtggacctgggaagcagcgtgcacagcggcctgATCGGCCCCCTGCTGGTGCCA	3019
Db	3598	cactaacacactgaacctgctcatgtgggagacaagtgcacagtacaggaatttgcctgt	3657
Qy	3020	caccAACACCCCTGAACCCCCCAGCGGAGGcagctCACTGTGCGGAATTTGCCCTGT	3079
Db	3658	ttcaccatctttgatgacccaaaagctggatctactgaaatatggagaagaaactg	3717
Qy	3080	CTTCAACATCTTCGACGAGACTTAAGAGCTGGTACTTCAACCGACACATGAGCGCACTG	3139
Db	3718	cagggtccctgcaatatccagatgaagatcccacttttaagagaattatcggtcca	3777
Qy	3140	cgccggcccttgcacATCAGATGGAATCCACCTCTCAGAGGAACTACCGCTTCCA	3199
Db	3778	tgcaatcaatggctacataatggatcacactacctggcttagtaatggctcagatcaag	3837
Qy	3200	cgccatCAACGGCTACATATGACACCCCTCGCCGCGCTGGTGAATGCCCCAGGACGCG	3259
Db	3838	gattcagatggctatcgtccagatgggcagcaatgaaacatccatctattcatttcag	3897
Qy	3260	CATCCGCTGGTACTCTGTCTATGGGAGCAAGAGACATCCACGATTCACCTTCAG	3319

Db	3898	tgacatgtgttcactctacgaataaagaggagtagataaaatggcactgtacaatatctcta	3957
Qy	3320	CGGCCAGGTTTTACCGTGCACGAAGAGGAGTCAAGATGGCCCTGTACAACCTGTA	3379
Db	3958	tcagggtgtttttgacacagtgaaatgttaccatccaaagctggaattttggcgggtgga	4017
Qy	3380	CCCCGGCGTGTTCGACGACTGTGGAGATGCTGCCCAACGAGCGCGGATCTGGCGCGTGA	3439
Db	4018	atgccttatctggcagcatctacatgctgggatgagcacactttttctggtgtacagcaa	4077
Qy	3440	GTGCTGTATCGCGGACGACCTTCGACGCGGATACGACCCCTGTTCTCTGGTACAGCA	3499
Db	4078	taagtgtcagactccccctgggaatggtctctggacacattagagattttcagattcacgc	4137
Qy	3500	CAAGTGCACGACCCCTGGGCATTCGACGGCGCACATCGCGCACTTCAGATCAACCG	3559
Db	4138	ttcaggacaatatggacagtggcccaaaagctggccagactcattattccgcatcaat	4197
Qy	3560	CAGCGGCAGTACGGCGAGTGGGCTCCCAAGCTGGCCGCCCTGCATCAGCGCGCAGCAT	3619
Db	4198	caatgctgtggagcaccagagaccctttcttgatcaaaagtggatctgttggcaccaat	4257
Qy	3620	CATGGCTGGTGCAGCCAGGAGCCCTTCTCTCGATCAAGGTGCACTGCTGGGCCCAT	3679
Db	4258	gattattcacggcatcaagaccagggtgcccttcagaagtttccagactctacatctc	4317
Qy	3680	GATCATCCACGGCATCAAGCCAGGGGGCGCCGCAAGTTACGACGCTGTATCATCG	3739
Db	4318	tcagtttatcatcatgattctgtgatgggaagaagtgccagacttatcaggaaattc	4377
Qy	3740	CCAGTTCATCATCTACTCTCTAGCGGCAAGATGGGAGACTACCGCGGCACAG	3799
Db	4378	caetggaaacctaatggtctcttttggcaatgtgattcatctgggataaaacaacatat	4437
Qy	3800	CACGGGACCCCTGATGGTGTCTTCGGCAAGTGGACAGCAGCGGGCATCAGCAACAAT	3859
Db	4438	ttttaacctccaattattcgtgatacatcgtttggacccaactcattatgactcg	4497
Qy	3860	CTTCAACGCCCCCATCATGCGCGCGCTACATCGCGCTGCACCCACCCACTACAGCATCGG	3919
Db	4498	cagcaactcttgcgatggagtgtgatggctgtgatttaaatagttgcagatgcatttggg	4557
Qy	3920	CAGCACCCCTGGGCATCGAGCTGATGGGCTTGGCACTCAACAGCTGCAGCATGCCCTGGG	3979
Db	4558	aatggagagtagaagaatatcagatgcacagattactgcttcactcacttaccataat	4617
Qy	3980	CATGAGAGCAAGGCCATCAGCGACGCCAGATCAACGGCTCCAGCTACTTCACCAACAT	4039
Db	4618	gittgcacctggtctcttccaaagctcagacttcactccaaaggaggagttaatgcctg	4677
Qy	4040	GTTCGCCACCTGGAGGCCCAACGAAGGCCGCCCTGCACCTGCAGGGCGGCGCAACGCCG	4099
Db	4678	gagacctcaggtgaataatccaaaagatggctgcaagtggacttcagaagacaatgaa	4737
Qy	4100	CGGCCCCCAGGTGAACMACCCCAAGGAGTGGCTGCAGTGGAGCTTCAGAAACCATGA	4159
Db	4738	agtcacaggagtaactactcagggagtaaaatctctgcttaccagcatgtatgtgaaga	4797
Qy	4160	GGTACTTGGCGTGACCCACAGGCGCTCAAGAGCCTGCTGACAGCATGTAGCTGAAGA	4219
Db	4798	gttccctcatctccagcagtcagaatggccatcagtgagactctcttttttcagaatggcaa	4857
Qy	4220	GTTCCTGTATCAGACGACGACGCGCCACGACGCGGACGAGTGGAGCTGTCTTCCAAACGCA	4279

Db	5994	tccagggtgtttttgagacagtgaaagttaaccatccaaagctggaaatttgccgggtgga	6053
Qy	3380	ccccggcgcttttcgacactgtggagatgctgccacgaagccgggatactggcgctgga	3439
Db	6054	atgccttattggcgagcatctacatgctgggatgagcacacttttcttggtgtacagcaa	6113
Qy	3440	gtgcctcatccggcgagacactgcacggcggcatgagcaccctgttcttggtgtacagcaa	3499
Db	6114	taagtgtcacagtccccctgggaatggctctctggacacattagagattttcagattacagc	6173
Qy	3500	caaagtgcacaccccccttgggcatggccagcggccacatccgggacattccagatcacgc	3559
Db	6174	ttcaggacaatatgacacdgtagcccccgaagctggccagacattcattatccagatcaat	6233
Qy	3560	cagcggccactgacggccagtgggctcccaagctggcccgctgcactacagcggccagcat	3619
Db	6234	caatgcctggagcaccaagagaccctttcttgatcaaggtggatctggttggcaccat	6293
Qy	3620	caaccccttgctcgaccaaagagacccttctcttggatcaaggtggaccttgcgtggccccc	3679
Db	6294	gattatcacggcatcaagacccagggtgcccgtcagaagttctccagctctcacatctc	6353
Qy	3680	gatcatccaggcatcaagacccaggcgccggccagagtttcagcagcctgtatcatcag	3739
Db	6354	tcagttatcatcatgtagttgatgtgggaagatggcagacattatcagagaaattc	6413
Qy	3740	ccagtttcattcatctacttcttagacggcgaagaatggcgacacctacggcgccacag	3799
Db	6414	cactggaaaccttaatggtctcttggcaatggattcatctggataaaacacaat	6473
Qy	3800	cacgggacacctgaatgggtttcttggcaacctggaacagcggcgcatcagacacacat	3859
Db	6474	ttttaacctccaattattgctcgatatacctggttgcaccaacctcattatagattcg	6533
Qy	3860	cttcaacccccccatcatcgcccgctgacatccgcctgcacccacccactacagcatcgc	3919
Db	6534	cagcactcttcgcatggattgatgggtgtgatttaaatagttgcacatgcattggg	6593
Qy	3920	cagcaccttcgcgatggagctgatggcgctgcacactgaacagctgcacatgcgccctggg	3979
Db	6594	aatggagagtaaagacaatatcagatgcacagatctgcttcaactcctacttaccat	6653
Qy	3980	catgcagacgaagcccatcagcgacgccacagatcacccgctccagctacttccaccaat	4039
Db	6654	gtttgccacctggtctcttccaaagctcgacttcaactccaaggaggagtaagtccgtg	6713
Qy	4040	gtttgccaccttggagccccagacgaagccgctgcacctgcagggcgccagcaacgcttg	4099
Db	6714	gagacctcagtgtaataatccaaagagtggtgcgaagtggactccagaagacaatgaa	6773
Qy	4100	gggccccaggtgcacacccccagagtggtgcaggtgcattccagaaacccatgaa	4159
Db	6774	agtccaggagtaactactcaggagtagaaatctctgcttaccagcatgtatggaagga	6833
Qy	4160	ggtgactggcgctgacacccagggcgctgaagagccctgctgaccacgactgtacgtgaagga	4219
Db	6834	gttctctcatctccagcagtcgaatggccatcagtggaactctcttttccagatggcaa	6893
Qy	4220	gtttctgactcagcagcgcaggaaggccaccagctggagaccctgttcttccaaaacggcaa	4279
Db	6894	agtaaaaggttttccagggaatatcaagactccttcacacctgtggtgaactctctagacc	6953
Qy	4280	ggtgaagtgctttccgggcacccagacagcttcaacccggtcgtgaacaccttgacac	4339

[illegible]

QY	2184	GGCATGACTGCCCTGCTGAAAGTCTCCAGCTGCGACAGCAACCGCGCACTACTACGAG	2243
Db	2327	gacagttatgaagataatttcagcatacttgcctgagtaaaacaatgccattgaaccaaga	2386
QY	2244	GACAGCTACGAGGAGATCTCGGCTACTGCTGCTCCAGAACACACGCCATCGAGCCCGCG	2303
Db	2387	agctctctccagaatttcagacacccgtagcactaggcaaaagcaatttaattgcccac	2443
QY	2304	TCTTTTCTCCAAAATCCCGCCACCCGACGCGCTCAGACGAGTTCAAGGCCACC	2360

Db	328	aacacctcagtcgtgtacacaaagactctgtttgtagaattcaaggttcaacttttcaac	387
Qy	204	AACACCAAGGCTGGTTACAAGAAACCCCTGTTCTGGAGTTCAACGACCACTGTTCAAC	263
Db	388	atcgtctaagcaagccacacctgtagtgggtctgtaggtctctaccatccaggtcaggtt	447
Qy	264	ATTGGCAAGCCGCCGCCCTCGATGGCCCTGCTGGGCCCCACCATCCAGAGCCGAGGTG	323
Db	448	tatgatacagtggttcattacaacttaagaacatggcttccactcctgtcagttctcatgct	507
Qy	324	TACGACACCGTGGTGATCACCTCGAAGAACATGGCCAGACCCCGCTCACCTGCACGCC	383
Db	508	gttggtgtatctactcgtgaagacttctgaggagcttgaatgatgatcagacaggtcaa	567
Qy	384	GTGGGCGTGAGCTACTTGGAGGGCCACGACGGGGCGGAGTACGACGACCAAGTCCGAC	443
Db	568	agggagaagaagatgataagctcttccctggttggaagccatacatatgtctgcgaagtc	627
Qy	444	CGCGAAGACGAGCACAAGGTTTCCCGGGGGGAGGCCACACTACTGCTGGCAGGTG	503
Db	628	ctgaaagagaagtgcataagtcgctctgaccacactgtgccttaectactacatatctttc	687
Qy	504	CTTAAGGAGAACGGCCCTATGGCCACGGACCCCTGTGGCTGACCTACAGTACCTGAGC	563
Db	688	catdgtgacctggttaaaagacttgaaatcaggccctcattggaagccctactagtagtaga	747
Qy	564	CACGTGGACCTGGTGAAGGATCTGAACAGCGGGGTGATCGGGGCCCTGCTGGTGTGCG	623
Db	748	gaaggagtcgtgcccaagaaaaagacacagaccttgcaacaaattataactacttttgc	807
Qy	624	GAGGGAGCTGGCCNAGGAGAAACCCAGACCTTGCACAAGTTTCCTCTGCTGTTCGCC	683
Db	808	gtatttgatgaagggaagaaagtggcaactcagaacaaagaacctcttgatgcagatagg	867
Qy	684	GTGTTTCAGCAGGAGGAGACGCTGGCACGCGAGACTGAACAGCGCTGATGCAGACCGC	743
Db	868	gatgtgcactctgctcgggcctggcctaataatgcacacagtcgaatggttatgtaaacag	927
Qy	744	GACGCGCCACGCGCCCGCGCTGGCCCAAGATGCACACCGTTAAGCGGCTACGTGAACCGC	803
Db	928	tctctgccaggtctgattggatgccacaggaataactcgtctattggcattgattggaatg	987
Qy	804	AGCCTGCGCGGCTGATCGGCTGCCACCGCAAGAGCGCTACTGGCACGTCATCGGCATG	863
Db	988	ggcaccactcctgaagtgacactcaatactctcgaaggtcacacaatttctgtgaggaac	1047
Qy	864	GGCACCACCCCTGAGGTGCACAGCATCTTCTCGAGGGCCACACTTCCTGCTGGTGGCAAC	923
Db	1048	catcgccagcgtctcttgaaatctcgccaaatacttctcactgctcaaacactcttg	1107
Qy	924	CACCGCCAGCGCCCTGGAGATCAGCCCCATCATCTCTGACTGCCACAGACCCCTGCTG	983
Db	1108	atggaccttgagcagtttctactgttttgtcatacttctcccaaccaactgatggcatg	1167
Qy	984	ATGAGACCTTAGGCCATCTTCTGTGTTCTTGCCACATCATGAGCGCCACGACGACGGCATG	1043
Db	1168	gaagcttatgtcaagtagacagctgtccagaggaacccccactacgaatgaaataaat	1227
Qy	1044	GAGGCTTAGCTGAGGTTGGACGCTGGCCCGGAGGACGCCACGCTGGCGATGAAGAACAC	1103
Db	1228	gaagaagcggagactatgatgatcttactgattctgaaatggatgtggtcaggttt	1287
Qy	1104	GAGGAGCCGAGGATACGACGACGACCTGACCGAGGAGGATGTCGTACGCTTC	1163

